Stochastic invertibility and related topics

Rémi Lassalle

2012-ENST-030

EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

présentée et soutenue publiquement par

Rémi LASSALLE

le 27 Juin 2012

Inversibilité stochastique

et thèmes afférents

Directeur de thèse : **Ali Süleyman ÜSTÜNEL**

Jury

Mme Hélène AIRAULT, Professeur, Université de Picardie	Examinateur
Mme Mireille CHALEYAT-MAUREL, Professeur, Université Paris VI	Examinateur
Mme Ana Bela CRUZEIRO, Professeur, Universidade de Lisboa	Rapporteur
M. Rama CONT, Directeur de Recherche, C.N.R.S.	Examinateur
M. Shizan FANG, Professeur, Université de Bourgogne	Président
M. Ivan NOURDIN, Professeur, Université de Loraine	Examinateur
M. Nicolas PRIVAULT, Professeur, Nanyang Technological University	Rapporteur
M. Ali Süleyman ÜSTÜNEL, Professeur, Télécom ParisTech	Directeur

TELECOM ParisTech
école de l'Institut Télécom - membre de ParisTech

T H È S E

Stochastic invertibility and related topics

Rémi Lassalle

Remerciements

Je tiens à exprimer ma profonde gratitude envers A.S. Üstünel pour avoir accepté de diriger mes recherches, ainsi que pour ses conseils de tous ordres, pour sa grande disponibilité, pour ses recommandations bibliographiques toujours profondes, ainsi que pour sa grande bienveillance. Je tiens également à remercier Ana Bela Cruzeiro et Nicolas Privault, qui ont rapporté sur ma thèse, pour le temps qu'ils y ont accordé, et pour leurs recommandations qui ont permis une nette amélioration du manuscrit dans la forme. Je remercie également le Jury dans son ensemble : c'est un honneur d'exposer mes résultats devant une telle assemblée. Plus généralement, je remercie l'ensemble des personnes qui ont permis le bon déroulement de cette thèse. Dans les aspects matériels, je remercie d'abord le directeur du département infres Gérard Memmi, et dans une moindre mesure son prédécesseur Michel Riguidel, lesquels ont permis de financer ce travail. Je remercie également le personnel de l'école doctorale qui a veillé au bon déroulement de la thèse dans ses aspects administratifs. Ici, je pense tout particulièrement à Henri Maître ainsi qu'à Florence Besnard. Dans les aspects logistiques, ma gratitude va à Patrick Clement qui a toujours été présent pour résoudre efficacement les divers problèmes pratiques rencontrés dans ces années de préparation. Je remercie également Jérôme Valentin et Paul Bourgade pour des références utiles au début de la thèse, ainsi que Marc Sigelle et Laurent Decreusefond pour leurs remarques stimulantes et leur vif enthousiasme, lorsqu'une partie substantielle de ces résultats leur ont été présentés en Juin de l'année précédente.

Finalement, je remercie ma famille dans son ensemble, mes amis, ainsi que tous ceux qui m'ont soutenu dans cette période.

Contents

Remerciements iii

Résumé vii

Introduction (in French) ix
- Une notion d'inversibilité stochastique sur l'espace de Wiener ix
- Structure de la thèse xiii
- Présentation des résultats principaux xiii

Abstract xv

Introduction xvii
- A notion of stochastic invertibility on Wiener space xvii
- Summary and main results xix

Part 1. Invertibility of adapted shifts associated with probabilities equivalent to the Wiener measure 1

I. **Preliminaries and notations** 3
1. Notations 3
2. The Girsanov shift 5

II. **Invertibility of adapted shifts on Wiener space** 9
1. Invertibility of adapted shifts 9
2. Invertibility and Entropy 11

III. **Local properties** 15
1. Local invertibility : definition and problem 15
2. Solution under finite energy condition 15
3. General case 16

IV. **Applications to stochastic mechanics: strong solutions for some Markovian equations** 19
1. A sufficient condition for the invertibility of Markovian shifts with states in $\mathbb{R}^d$ 19
2. Invertibility of free Schrödinger shifts 20
3. Invertibility of the free Schrödinger shifts in the perspective of optimal transport 26

V. **Applications to information theory : Information loss on the path space and Shannon's inequality** 31
1. Talagrand's inequality and Sobolev's inequality on any abstract Wiener space 31
2. Information loss on the path space 34
3. Üstünel's criterion in terms of variance and Shannon's inequality 38

Part 2. A notion of stochastic invertibility on Wiener space 43

VI. **Invertibility of Girsanov shifts on abstract Wiener space : general case** 45
1. Introduction 45
2. Preliminaries and notation 46
3. The Girsanov shift on abstract Wiener space 52
4. Pull-back of the Girsanov shift with related morphisms 56
5. First applications, the innovation process 59
6. Entropy, energy, and mean square error 62
7. Variational formulation of the entropy 63
8. Local properties 68
9. The Monge problem and the invertibility 69
10. Application to the innovation conjecture 71

VII. **Stochastic invertibility on Wiener space for stochastic differential equations with dispersion** 73
1. Introduction 73
2. Preliminaries and notations 74
3. The Brownian transform 76
4. Invertibility of the Brownian transform 78
5. The structure of quasi-invariant flows 81
6. Remark : weak derivative along the quasi-invariant flows 82
7. Local properties 83
8. Criterion of invertibility 84
9. Some further extensions 85

Part 3. Towards stochastic differential Geometry 87

VIII. **Invertibility of adapted shifts on the Lie-Wiener space** 89
1. Introduction 89
2. Preliminaries and notation 90
3. Transformations of measure induced by adapted shifts on Lie groups 94
4. The innovation process 99
5. Invertibility 101
6. Criterion of invertibility based on the relative entropy 102

Bibliography 105

Résumé

Un morphisme d'espaces de probabilité vers l'espace de Wiener, qui est de plus adapté, peut être associé canoniquement à certaines lois d'équations différentielles stochastiques, une de ses propriétés essentielles étant qu'il est un isomorphisme d'espaces de probabilité si et seulement si l'équation différentielle stochastique associée admet une unique solution forte. Puisqu'on peut le voir comme un mouvement Brownien canoniquement associé à une loi, on lui donnera le nom de transformée Brownienne de la loi, et il s'agira de s'intéresser à son inversibilité. Comme on le verra en la présentant, cette notion, qui s'enracine autour de résultats déjà employés par Föllmer il y a fort longtemps, prolonge naturellement la notion d'inversibilité des dérives adaptées qui a été développée ces dernières années dans les travaux d' Üstünel et de Zakai où elle se trouve déjà en germes. Cette approche présente l' avantage d'impliquer des notations compactes, ainsi que de permettre une exploitation directe des puissants théorèmes de l'analyse fonctionnelle, ce qui produit naturellement des énoncés très généraux, et des preuves d'une grande efficacité. Par ailleurs du fait même qu'elle s'exprime en termes d'isomorphismes d'espaces de probabilité, cette notion se transporte naturellement dans des cadres géométriques. Pour exemple on la transportera dans l'espace des chemins à valeurs dans des groupes de Lie de dimension finie. De plus, de nombreux problèmes issus de thèmes très divers s'intriquent spontanément avec cette notion, ce qui ouvre un large spectre d'applications et permet d'aborder à la fois des problèmes fondamentaux, mais aussi des problèmes très concrets et très appliqués. Dans le cas de probabilités absolument continues par rapport à la mesure de Wiener, on sera naturellement amené à envisager des problèmes issus de la théorie du filtrage, de la physique statistique, du contrôle stochastique, du transport optimal, ainsi que la théorie de l'information. En particulier, on donnera un résultat d'unicité trajectorielle très général pour la représentation stochastique de la mécanique quantique en temps Euclidien, et on étendra l'inégalité de Shannon aux espaces de Wiener abstraits, cette dernière recevant au passage une jolie interprétation en termes de perte d'information dans un canal Gaussien.

Introduction (in French)

La structure de cette introduction est la suivante. Dans un premier temps je préciserai dans quelle acception on entend ici le terme d'inversibilité. Je rappellerai ensuite le cadre précis dans lequel cette notion a été progressivement introduite dans les travaux antérieur à cette thèse, où elle consiste essentiellement en l'inversibilité de certaines dérives, anticipatives et non-anticipatives sur l'espace de Wiener. J'esquisserai ensuite le mouvement général de cette thèse en montrant comment cette notion s'élabore. En particulier j'indiquerai en quoi les concepts développés ici s'appuient sur les travaux antérieurs, et en quoi ils s'en distinguent : précisément je montrerai comment on se détache progressivement d'une notion d'inversibilité attachée à certaines perturbations de l'identité adaptées arbitraires, pour s'intéresser à l'inversibilité de certains morphismes d'espaces de probabilité canoniquement associés à certaines lois d'équations différentielles stochastique, celle-ci généralisant et éclairant celle là. Je présenterai ainsi, en la justifiant, la structure générale de cette thèse ainsi que les thématiques abordées. Enfin, je donnerai un résumé détaillé où j'indiquerai les résultats obtenus qui me paraissent les plus significatifs.

Une notion d'inversibilité stochastique sur l'espace de Wiener

Commençons par expliquer la manière dont on emploiera généralement le terme d'inversibilité. Soit W un espace de Banach séparable et $\mathcal{P}(W)$ des probabilités boréliennes sur W. Pour $\nu, \eta \in \mathcal{P}(W)$ considérons deux applications mesurables $V : W \to W$ et $U : W \to W$. Si les lois images $V\nu$ et $U\eta$ vérifient $V\nu << \eta$ (i.e. absolument continue) et $U\eta << \nu$, alors la composition $V \circ U$ (resp. $U \circ V$) est définie sans ambiguité comme classe d'équivalence par rapport à η (resp. par rapport à ν). On peut ainsi très naturellement définir une notion d'inversibilité par rapport à η et à ν de la manière suivante : soit $V : W \to W$ tel que $V\nu << \eta$ on dira que V est inversible par rapport à η et à ν si on peut trouver un $U : W \to W$ dont la loi image vérifie la condition d'absolue continuité $U\eta << \nu$ et tel que $\eta - p.s.$

$$V \circ U = I_W \tag{1}$$

et $\nu - p.s.$

$$U \circ V = I_W \tag{2}$$

Dans (1) (resp. dans (2)) l'égalité est donnée $\eta - p.s.$ (resp. $\nu - p.s.$) car $V \circ U$ (resp. $U \circ V$) est défini $\eta - p.s.$ (resp. $\nu - p.s.$), et $I_W : \omega \in W \to \omega \in W$ désigne l'identité sur W ce qui justifie l'emploie du terme d'inversibilité. Par la suite W désignera l'espace de Wiener classique, H son espace de Cameron-Martin associé (voir par exemple **[25]**), et μ la mesure de Wiener : c'est dans ce cadre, ou dans certaines de ses généralisations que l'on travaillera. Rappelons que l'espace de Wiener classique $W = C([0, 1], \mathbb{R}^d)$ est l'ensemble des fonctions continues à valeurs dans $\mathbb{R}^d$. Un élément $\omega \in W$ représente une trajectoire et on notera $t \to W_t(\omega) := \omega(t)$ le processus coordonnées. La mesure de Wiener μ est telle que lorsqu'on équipe W de cette mesure, le processus coordonné est un mouvement Brownien pour sa propre filtration. L'espace de Cameron-Martin associé est défini par

$$H := \left\{ h : t \in [0, 1] \to h_t := \int_0^t \dot{h}_s ds : \int_0^1 |\dot{h}_s|^2_{\mathbb{R}^d} ds < \infty \right\}$$

et sa propriété essentielle est que les translations selon H préservent les classes d'équivalences par rapport à la mesure de Wiener. Plus précisemment pour tout $h \in H$, si on note $\tau_h : \omega \in W \to \tau_h(\omega) := \omega + h \in W$, alors la mesure image $\tau_h\mu$ est équivalente à μ. Afin de mieux faire ressortir les spécificités de cette thèse, replaçons maintenant les choses dans leur contexte historique. La notion d'inversibilité de transformations sur l'espace de Wiener a été introduite puis intensivement développée par Üstünel et Zakai dans des contextes variés ces dix dernières années. Afin de parler précisément, introduisons d'abord des notations concernant les perturbations (adaptées) de l'identité. Pour $\nu \in \mathcal{P}(W)$ on notera $L^0(\nu, H)$ l'ensemble des classes d'équivalences obtenues en identifiant les $u : W \to H$ mesurables qui coincident ν presque partout. En particulier on s'intéressera à $L^0_a(\nu, H)$ qui est le sous ensemble de $L^0(\nu, H)$ constitué par les $u \in L^0(\nu, H)$ tel que $(\dot{u}_s, s \in [0,1])$ soit adapté à la filtration engendré par le processus coordonné qu'on note $t \to W_t$: quand on parlera de processus adaptés ce sera par rapport à cette filtration. Soit $\nu \in \mathcal{P}(W)$ et $V := I_W + v$ où $v \in L^0(\nu, H)$ et où $I_W : \omega \in W \to \omega \in W$: V est appelée une perturbation de l'identité, ou encore une dérive. Il s'agit, dans ces travaux antérieur de partir d'une perturbation de l'identité V telle que $V\mu << \mu$ et de chercher à l'inverser : c'est ainsi que s'est développé le thème de l'inversibilité des perturbations de l'identité. De nombreux résultats ont ainsi été mis en évidence. Dans le cas où V est non adapté, l'inversibilité intervient entre autre dans le théorème de Ramer (voir **[45]**), et en transport optimal puisque l'optimum du problème de Monge est inversible (voir **[12]**,**[13]**). Si l'inversibilité de certaines dérives anticipatives interviendra parfois dans cette thèse, pour l'essentiel nous nous intéresserons à l'inversibilité de transformations adaptées. L'étude de l'inversibilité de transformations adaptées sur l'espace de Wiener a été entreprise plus récemment par Üstünel et Zakai sous des hypothèses bien précises, et ce sont ces travaux (i.e. **[53]**,**[46]**,**[47]**) qui constituent à proprement parler le point de départ de ce travail. L'inversibilité des dérives adaptés était défini essentiellement par deux hypothèses vraisemblablement héritées en partie de leurs travaux dans le cas non adapté. La première qu'on notera $(H-1)$ est simplement

$$v \in L^0_a(\mu, H)$$

(a pour adapté), et la seconde qu'on notera $(H-2)$ porte sur l'intégrabilité exponentielle

$$E_\mu \left[\exp\left(-\int_0^1 \dot{v}_s dW_s - \frac{1}{2}\int_0^1 |\dot{v}_s|^2_{\mathbb{R}^d} ds \right) \right] = 1 \tag{3}$$

Dans ce contexte, ils définissent l'inversibilité ainsi : on dira qu'un $V := I_W + v$ où v satisfait $(H-1)$ et $(H-2)$ est inversible si on peut trouver un $U : W \to W$ mesurable tel que $U\mu << \mu$ et $\mu - p.s.$

$$V \circ U = I_W \tag{4}$$

et $\mu - p.s.$

$$U \circ V = I_W \tag{5}$$

Dans ce cas précis il s'agissait donc exclusivement de classes d'équivalence pour la mesure de Wiener. Certaines conditions suffisantes d'inversibilité pour ces dérives adaptées sous $(H-1)$ et $(H-2)$ ont ainsi été mises en évidence, essentiellement au moyen du calcul de Malliavin et plus généralement de l'analyse stochastique, tout comme c'est évidemment le cas concernant leurs travaux sur l'inversibilité des dérives anticipatives. Deux motivations ont été avancées (dans **[53]**) pour justifier de la pertinence de l'inversibilité des perturbations adaptées de l'identité. La première, qui trouve en fait son origine dans des problématiques connexes au théorème de Ramer porte sur la densité. Etant donnée un shift $U := I_W + u$, $u \in L^0_a(\mu, H)$, si U possède un inverse V, alors on dispose de la formule

$$\frac{dU\mu}{d\mu} = \exp\left(\int_0^1 \dot{u}_s dW_s + \frac{1}{2}\int_0^1 |\dot{u}_s|^2_{\mathbb{R}^d} ds \right) \circ V$$

La seconde consiste en un lien avec les équations différentielles stochastique : l'inversibilité de $V := I_W + v$ équivaut à l'existence d'une unique solution forte pour

$$dX_t = dB_t - \dot{v}_t \circ U dt; X_0 = 0 \tag{6}$$

Cette seconde motivation s'avèrera tout à fait essentielle dans la seconde partie de la thèse puisqu'on s'appuiera dessus pour généraliser la notion d'inversibilité des perturbations adaptées de l'identité : il ne s'agira plus d'inverser des perturbations de l'identité arbitraire, mais d'inverser certaines transformations associées à des lois de solutions d'équations différentielles stochastiques. Mais avant d'aborder ce point, rappelons un résultat récent qui justifie la connexion avec de nombreux thèmes envisagés dans cette thèse. Suite à ces recherche de conditions suffisantes d'inversibilité, ([**53**]) Üstünel a également mis en évidence (dans [**47**]) une condition nécessaire d'inversibilité pour les dérives adaptés sur l'espace de Wiener (généralement non- Markoviennes[1]) : il s'agit d'un critère saisissant qui se fonde sur une égalité entre énergie cinétique et entropie. On rappelle que pour deux probabilités η et ν telles que $\nu << \eta$ (i.e. absolument continue), leur entropie relative est définie par

$$H(\nu|\eta) := E_\nu \left[\ln \frac{d\nu}{d\eta} \right]$$

Pour être concret, soit $V := I_W + v$ (resp. $U := I_W + u$) induit par un $v \in L_a^0(\mu, H)$ (resp. un $u \in L_a^2(\mu, H)$) qui satisfont tous deux (3). Si U et V sont de plus liés par la relation

$$\frac{dU\mu}{d\mu} = \exp\left(-\int_0^1 \dot{v}_s dW_s - \frac{1}{2}\int_0^1 |\dot{v}_s|_{\mathbb{R}^d}^2 ds \right)$$

alors on a

$$2H(U\mu|\mu) \leq E_\mu \left[|u|_H^2 \right]$$

avec égalité si et seulement si V et U sont inverse l'un de l'autre. L'objectif de cette thèse était initialement de généraliser ce critère, d'en développer des applications dans de nombreux domaines, et de contribuer aux développement de l'étude de l'inversibilité des perturbations de l'identité adaptées sur l'espace de Wiener. Ces développements font l'objet de la première partie de ce manuscrit qui s'étend des chapitres I à V. La seconde partie de cette thèse (Chapitre VI-VII), consiste à retourner le problème dans l'autre sens, ce qui aboutit à une notion d'inversibilité stochastique sur l'espace de Wiener. Il ne s'agira plus de partir d'une dérive de l'identité arbitraire et de chercher des conditions suffisantes à son inversibilité, puis de justifier de sa pertinence en évoquant un lien avec une équation différentielle stochastique. Au contraire, il s'agira de partir d'une loi d'équation différentielle stochastique, de lui associer une certaine transformation et de chercher à l'inverser. Donnons avec plus de détails les différentes étapes qui conduisent à cette notion. Il s'est d'abord agit de généraliser la notion d'inversibilité de dérives adaptées qui induisent une loi ν absolument continue par rapport à la mesure de Wiener, mais pas équivalente. On s'appuie alors sur le résultat élémentaire suivant dont l'usage est suggéré par les hypothèses du critère d'inversibilité fondé sur l'entropie. On rappelle en effet le résultat suivant, déjà employé par Föllmer dans les annés soixante-dix, mais qu'on peut lire aussi implicitement dans certains résultats de mécanique stochastique datant de la même époque : étant donné une probabilité ν absolument continue par rapport à la mesure de Wiener, on peut trouver un unique $v \in L_a^0(\nu, H)$ tel que

$$\frac{d\nu}{d\mu} = \exp\left(-\int_0^1 \dot{v}_s dW_s - \frac{1}{2}\int_0^1 |\dot{v}_s|_{\mathbb{R}^d}^2 ds \right) \tag{7}$$

[1]Dans le cas Markovien, ce résultat semble en partie connu depuis les années 90, où une partie du résultat apparait sous des hypothèses restrictives dans des travaux en lien avec la mécanique stochastique, voir par exemple [**34**] (le Théorème 1.2, et les références qui s'y trouvent). Néanmoins, on est alors dans un cadre Markovien, et il est n'est pas alors réellement question du problème d'existence d'une unique solution forte qui est le propre de l'inversibilité. Notons encore que nous avons appris récemment le lien de ce résultat avec le résultat principal de [**3**] : le lien entre ces formules n'est pas envisagé dans cette thèse.

Dans cette thèse on appellera $V := I_W + v$ la dérive de Girsanov associée à ν ce qui rappelle bien l'origine de cette dérive. Rappelons qu'une application directe du théorème de Girsanov prouve que $t \to V_t$ est Brownien sur $(W, \mathcal{F}^{\nu}_{\cdot}, \nu)$. Ainsi V est un morphisme de probabilité (adapté) entre $(\mathrm{W}, \mathcal{F}^{\nu}, \mu)$ et $(\mathrm{W}, \mathcal{F}^{\mu}, \mu)$, canoniquement associé à ν. On dira naturellement que V est $\nu-$inversible si il existe un $U : W \to W$ mesurable et adapté tel que $U\mu << \nu$ et $\mu - p.s.$

$$V \circ U = I_W \tag{8}$$

$\nu - p.s.$

$$U \circ V = I_W \tag{9}$$

L'inversibilité de V est donc précisément la condition qui fait du morphisme V un isomorphisme d'espace de probabilités (filtrés). De même que dans le cas équivalent, l'inversibilité de la dérive de Girsanov associée à ν est équivalente à l'existence d'une unique solution forte pour (6) : la différence est que la loi de cette solution est maintenant absolument continue par rapport à μ et pas nécessairement équivalente. En outre, il est très facile de voir que l'étude de l'inversibilité des dérives de Girsanov associée à des lois $\nu << \mu$ généralise l'inversibilité des dérives adaptées telle qu'elle a été envisagée par Zakai et Üstünel, c'est à dire sous $(H-1)$ et $(H-2)$. En effet étant donné un v qui satisfait $(H-1)$ et $(H-2)$, alors $V = I_W + v$ est la dérive de Girsanov associée à la mesure ν équivalente à μ de densité donnée $\mu - p.s.$ par (7). Dans ce cas $\nu \sim \mu$ et les conditions (8) et (9) deviennent précisement (4) et (5). Néanmoins il ne s'agit généralement plus d'inverser n'importe quel $V := I_W + v$ où $v \in L^0_a(\nu, H)$, mais seulement les dérives de Girsanov associées à des lois $\nu << \mu$. La question se pose alors de savoir si l'on peut traiter de la même manière des lois d'équations différentielles stochastiques avec dispersion, dont la loi n'est plus absolument continue par rapport à la mesure de Wiener. La densité $\frac{d\nu}{d\mu}$ n'étant plus défini dans ce cas, il faut reformuler le résultat élémentaire qu'on a rappelé sur les dérives de Girsanov dans le cas $\nu << \mu$. Le fait d'être un morphisme adapté entre $(W, \mathcal{F}^{\nu}, \nu)$ et l'espace de Wiener ne caractérisant V qu'à une rotation près sur $(W, \mathcal{F}^{\nu}, \nu)$, il convient de se donner une condition plus précise. Or pour $\nu << \mu$ la dérive de Girsanov V est aussi caractérisé par le fait que (I_W, V) est solution de (6) sur $(W, \mathcal{F}_{\cdot}, \nu)$. Dans le cas d'une solution ν d'équation avec dispersion de la forme

$$dX_t = \alpha_t(X)dB_t + \beta_t(X)dt; X_0 = 0 \tag{10}$$

on peut encore, sous certaines conditions de non-dégénerescence du coefficient de dispersion, trouver un morphisme V canoniquement associée à ν, telle que (I_W, V) est solution de l'équation sur $(W, \mathcal{F}^{\nu}, \nu)$. Quand V existe, on l'appellera la transformée Brownienne de ν. Par définition, en tant que processus $t \to V_t$ est Brownien sur $(W, \mathcal{F}^{\nu}_{\cdot}, \nu)$, et en tant qu'application $V : W \to W$ est un morphisme d'espaces de probabilité entre (W, ν) et (W, μ) (V Brownien signifie $V\nu = \mu$). On définit alors la $\nu-$inversibilité de la transformée Brownienne V de ν de la même manière que précédemment comme l'existence d'un $U : W \to W$ mesurable et adapté tel que $U\mu << \nu$ et dont les compositions avec V satisfont (4) et (5). Dans ce contexte, hormis dans le cas d'une probabilité $\nu << \mu$ où la transformée Brownienne existe toujours et se confond avec la dérive de Girsanov, V n'est plus une perturbation de l'identité. Modulo des hypothèse de non dégénerescences, on peut encore montrer que (10) admet une unique solution forte si et seulement si la loi en question admet une transformée Brownienne inversible. Conceptuellement, on s'est ainsi totalement détaché l'idée d' inverser une perturbation de l'identité arbitraire, pour s'intéresser au problème d'existence d'un isomorphisme d'espace de probabilité, canoniquement associée à une loi ν d'équation différentielle stochastique, entre $(W, \mathcal{F}^{\nu}, \nu)$ et l'espace de Wiener. A travers cette notion, nous seront amenés à visiter et revisiter des champs divers, dont certains sont très récents, et d'autres sont déjà forts anciens. Dans cette thèse, outre des résultats nouveaux qui ont un intérêt propre, et des généralisations de résultats existants qui nous paraissent utiles ou éclairantes, on trouvera également de nouvelles preuves courtes de résultats connus. La présence de ces dernière, dans ce manuscrit, a pour but de convaincre le lecteur, comme nous l'avons été, de la puissance de ce cadre de travail et de la pertinence de cette notion. Trois publications sont pour le moment associées à cette thèse ([**27**], [**28**], [**29**]).

Structure de la thèse

Donnons à présent la structure de cette thèse qu'on vient de motiver : La partie I (Chapitre $I - V$), s'intéresse à l'inversibilité de dérives adaptées dans le cadre de travail d'Üstünel et Zakai i.e. sous $H-1$ et $H-2$. On y développe certains résultats fondamentaux et des applications à la mécanique quantique et à la théorie de l'information. En particulier, on donnera un résultat d'unicité trajectorielle pour la mécanique quantique en temps Euclidien, et on donnera une version de l'inégalité de Shannon pour les espaces de Wiener abstraits. Afin de faciliter la lecture en allégeant les notations, et en évitant certaines subtilités liées aux classes d'équivalences, on a préféré présenter cette partie sous $(H-1)$ et $(H-2)$. Néanmoins on indiquera en partie II comment ces résultats se généralisent à des mesures absolument continues. La partie II (chapitre VI-VII) définit la notion d'inversibilité stochastique qu'on a présenté plus haut i.e. la $\nu-$inversibilité d'une certaine transformée Brownienne canoniquement associée à une loi ν de solution d'E.D.S : on sort ainsi du cadre de travail d'Üstünel et Zakai. On traite d'abord le cas $\nu << \mu$ dans un cadre abstrait en donnant de nombreuses applications. Le second chapitre de cette partie traite le cas d'équations différentielles avec dispersion. Le dernier chapitre de la thèse est une ouverture vers la géométrie différentielle stochastique : l'essentiel du travail consiste à transférer la notion d'inversibilité des perturbations de l'identité à l'espace des chemins à valeurs dans certains groupes de Lie.

Présentation des résultats principaux

Partie I : Inversibilité des perturbations de l'identité associées à des probabilités équivalentes à la mesure de Wiener.

Chapitre I : Préliminaires et notations. On fixe le cadre de travail pour la partie I qui comprend les Chapitres I à V. On remarque également (Proposition I.1) qu'un changement de mesure permet de penser à la variance de la dérive de Girsanov comme à une entropie, ce qui sera très utile dans les applications à la théorie de l'information.

Chapitre II : Inversibilité, Critère d'inversibilité basé sur l'entropie. On commence par une généralisation (Théorème II.1) du critère d'inversibilité basé sur l'entropie qui relaxe les hypothèses sur le supposé inverse $U := I_W + u$. La preuve donnée qui semble nouvelle, est très courte et n'utilise ni la projection duale prévisible, ni ne processus d'innovation. Notons que celle-ci s'étend au cas de mesures absolument continues comme celà sera expliqué en Partie II. Ce résultat a de belles implications dans le domaine de la théorie de l'information qui sont développées au Chapitre V, mais aussi en physique statistique. On donne aussi des résultats de convergence. On montre également en toute généralité l'existence d'une expression explicite de l'inverse d'une dérive inversible arrêtée par un temps optionnel (Proposition II.2).

Chapitre III : L'inversibilité est une propriété locale. Dans cette partie on prouve qu'en toute généralité, l'inversibilité est une propriété locale au sens usuel du calcul stochastique (Théorème III.2). La preuve s'appuie sur les résultats du chapitre précédent. J'avais d'abord obtenu le résultat sous une condition d'entropie finie à l'aide du critère d'inversibilité fondé sur l'entropie. Ce résultat est le Théorème III.1.

Chapitre IV : Résultats d'unicité pour certains ponts de Schrödinger, lien avec le transport optimal. On prouve l'unicité trajectorielle pour la représentation stochastique de la mécanique quantique en temps euclidien dans le cas libre : en toute généralité en dimension fini (Théorème IV.2), et pour les densités bornés en dimension infinie (Théorème IV.3). La preuve s'appuie sur les propriétés locales de l'inversibilité qui permettent de localiser des résultats déjà existants en dimension finie, et sur un résultat obtenue par un théorème de point fixe et le calcul de Malliavin en dimension infinie. Un intérêt de ce résultat est de donner plus de consistance au modèle stochastique, puisque l'inversibilité implique l'unicité trajectorielle qui est le pendant stochastique du déterminisme des phénomènes physiques sous-jacents. Ce résultat permet d'obtenir une représentation variationnelle de l'entropie (Proposition IV.3). Une formule similaire est ensuite obtenue en transport optimal (Proposition IV.4).

Chapitre V : Application à certaines inégalités de la théorie de l'information. Le résultat principal de ce chapitre est l'extension de l'inégalité de Shannon aux espaces de Wiener abstraits (Théorème V.3). Celle-ci s'obtient à partir d'un changement de mesure qui permet de réécrire le critère d'inversibilité en terme de variance. A travers la preuve produite, l'inégalité de Shannon résulte de la propriété d'additivité de la variance, et de la perte d'information dans un canal Gaussien.

Partie II : Une notion d'inversibilité stochastique.

Chapitre VI : Inversibilité des dérives de Girsanov dans les espaces de Wiener abstraits. On se place dans le cadre très général d'un espace de Wiener abstrait, où une structure temporelle peut être obtenue à l'aide de certaines suites de projections sur l'espace de Cameron-Martin introduites dans [**49**]. Par rapport à l'espace de Wiener classique, la structure temporelle est apauvrie. Néanmoins ce cadre est suffisant pour construire le calcul stochastique. L'intérêt de ce cadre, outre sa plus grande généralité qui ouvre plus d'applications, est d'épurer les démonstrations en interdisant certaines manipulations. On démontre d'abord l'existence d'une dérive de Girsanov en toute généralité (Théorème VI.3). Une partie de ce théorème apparait déjà dans [**50**], mais avec une preuve fausse. On étudie ensuite l'inversibilité de ces dérives, généralisant ainsi les résultats de la première partie. On généralise alors les résultats de [**47**] sur la conjecture de l'innovation de la théorie du filtrage : on prouve qu'elle ne dépend que de la loi du signal reçu à travers l'inversibilité de la dérive de Girsanov associé (Thérorème VI.9). Un lien nouveau est aussi donné avec le transport de Monge (Thérorème VI.8 et son corollaire).

Chapitre VII : Inveribilité stochastique pour des lois d'équations différentielles stochastiques avec dispersion. On généralise la notion d'inversibilité à des équations du type

$$dX_t = \alpha_t(X)dW_t + \beta_t(X)dt; X_0 = x \tag{11}$$

Sous certaines hypothèses de non dégénerescence, sous lesquelles (8) présente des propriétés similaires au cas Brownien, une application $V : W \to W$ est associée à la loi ν d'une solution de (8). V étant un mouvement Brownien sur (W, ν) on l'appelle la transformée Brownienne de ν : dans le cas où $\nu << \mu$ V est la dérive de Girsanov. On généralise alors le critère d'inversibilité basé sur l'entropie dans ce cadre, et on prouve que l'inversibilité de V est reliée à l'existence d'une unique solution forte pour (8). D'autres extensions de cette notions sont également envisagées

Ouverture à la géométrie différentielle stochastique :

Chapitre VIII : Inversibilité sur l'espace des chemins à valeur dans un groupe de Lie. On utilise une exponentielle stochastique pour relever la mesure de Wiener sur l'espace des chemins à valeur dans certaines algèbres de Lie sur l'espace des chemins à valeur dans le groupe de Lie associé. On généralise ensuite dans ce contexte plusieurs résultats importants, dont le principal semble être le critère basé sur l'entropie (Théorème VIII.5)). On montre également que l'inversibilité de ces perturbations de l'identité est encore relié à l'existence d'une unique solution forte pour des équations différentielles stochastiques.

Abstract

In this work we investigate a notion of stochastic invertibility on Wiener space. Roughly speaking a morphism of probability spaces with values on the Wiener space, which is further adapted, can be canonically associated to the laws of the solutions to some stochastic differential equations. One of the main properties of this morphism is to be invertible (i.e. to be an isomorphism of probability spaces) if and only if the underlying stochastic differential equation has a unique strong solution. Since it may be seen as a Brownian motion, we call it the Brownian transform of the associated law, and we will study the invertibility of this Brownian transform. We will see that this notion, whose origins may be found in earlier results related to stochastic mechanics, extends and enlightens the notion of invertibility of adapted shifts on Wiener space which was investigated by Üstünel and Zakai in their recent papers, where this notion already appears clearly between the lines. Among other approaches to deal with strong solutions of stochastic differential equations, this approach takes advantage of the compact notations it involves, and of the powerful theorems of functional analysis which applies very naturally within this framework. Hence, this approach is very efficient : it provides very general results together with very short proofs. On the other hand, since it essentially uses some adapted isomorphisms of probability spaces, this notion naturally fits in more geometrical frameworks and is fully compatible with stochastic differential geometry. To illustrate this point, we will transport this notion on the space of the paths with values in a finite dimensional Lie group. Moreover, many problems arising in various fields are deeply related to this notion. This opens to a wide spectrum of applications, some of them being very concrete and applied. In the case of probabilities absolutely continuous with respect to the Wiener measure, we will investigate problems of various origins such as statistical physics, information theory, filtering, but also stochastic control and optimal transport. For instance, we will prove a very general result of pathwise uniqueness for the stochastic picture of euclidean quantum mechanics, and we will extend Shannon's inequality to any abstract Wiener spaces. This latter also receives a nice interpretation in terms of information loss in a Gaussian channel.

Introduction

A notion of stochastic invertibility on Wiener space

Roughly speaking, the notion of stochastic invertiblity on Wiener space investigated in this work amounts to express the problems of existence of a unique strong solution to a stochastics differential equation in terms of invertibility of morphisms of probability spaces, which are further adapted. We now explain formally the basic ideas. To be precise, we first introduce the set $W := C([0,1], \mathbb{R}^d)$ of the continuous paths with values in $\mathbb{R}^d$. For any Borelian probability ν on W we note $t \to W_t$ the coordinate process i.e. for any $\omega \in W$ and $t \in [0,1]$ we have $W_t(\omega) := \omega(t)$. This coordinate process generates a filtration $\left(\mathcal{F}_t^0\right)$ whose usual augmentation with respect to ν is noted $(\mathcal{F}_t^\nu)$. We recall that the Wiener measure μ is such that the coordinate process is a Brownian motion on (W, μ), with respect to its own filtration. This framework being fixed, we consider the stochastic differential equation

$$dX_t = \alpha_t(X)dB_t + \beta_t(X)dt; X_0 = 0 \tag{1}$$

where X_t takes values in $\mathbb{R}^d$, and where (B_t) is a d dimensional Brownian motion. Let us assume that this equation has a weak solution whose law is noted ν. Further assume that that there is a unique V such that (I_W, V) is a weak solution of (1) on $(W, \mathcal{F}^\nu, \nu)$. Under this assumption, as a process $t \to V_t(\omega) \in \mathbb{R}^d$ is a Brownian motion on $(W, \mathcal{F}^\nu_\cdot, \nu)$, and as a mapping $V : W \to W$ defines an adapted morphism of probability spaces between $(W, \mathcal{F}^\nu, \nu)$ and $(W, \mathcal{F}^\mu, \mu)$ (V is Brownian exactly means $V\nu = \mu$). We call $V : W \to W$ the Brownian transform of ν. Since $V\nu = \mu$, it is straightforward to check that for any $U : W \to W$ which is measurable whith $U\mu << \nu$, the pullback $V \circ U : W \to W$ (resp $U \circ V : W \to W$) is well defined $\mu - a.s.$ (resp. $\nu - a.s.$). Hence we can define V to be $\nu-$invertible, if there is a $U : W \to W$ measurable and adapted such that $U\mu << \nu$ and with the further property that $\mu-$a.s.

$$V \circ U = I_W \tag{2}$$

and $\nu-$a.s.

$$U \circ V = I_W \tag{3}$$

where $I_W : \omega \in W \to \omega \in W$ is the identity on the path space W. The interesting point is that, under some further non-degeneracy assumption, one can prove that V is invertible if and only if (1) has a unique strong solution. Hence, it is possible to turn the question of existence of a strong solution to (1) into the following : is V $\nu-$invertible ? Or is V an isomorphism of (filtered) probability space between $(W, \mathcal{F}^\nu, \nu)$ and the Wiener space ? In this thesis we will be mainly interested by the case where ν is a probability absolutely continuous with respect to the Wiener measure μ. In that case (as we will recall it below), ν is the solution of a stochastic differential equation of the shape

$$dX_t = dB_t - \dot{v}_t(X)dt; X_0 = 0 \tag{4}$$

where $\nu-$a.s.

$$\int_0^1 |\dot{v}_s|^2_{\mathbb{R}^d} ds < \infty \tag{5}$$

Moreover, the Girsanov theorem then yields that $(I_W, I_W + v)$ a weak solution of (4) on $(W, \mathcal{F}^{\nu}_{\cdot}, \nu)$. Hence, in that case the Brownian transform V of ν is given by

$$V = I_W + v$$

Since in that case the existence of V is a consequence of the Girsanov theorem, we also call it the Girsanov shift associated with ν.

In **[53]**, Üstünel and Zakai investigated the following notion, probably inherited from their study of Ramer's theorem (see **[45]** or **[50]**). Let $v := \int_0^{\cdot} \dot{v}_s ds$ be an adapted shift on $(W, \mathcal{F}^{\mu}_{\cdot}, \mu)$ which satisfies the hypothesis

$$E_\mu\left[\exp\left(-\int_0^1 \dot{v}_s dW_s - \frac{1}{2}\int_0^1 |\dot{v}_s|^2_{\mathbb{R}^d} ds\right)\right] = 1 \tag{6}$$

and which is such that $\mu - a.s.$ (5) holds. They consider the invertibility of $V := I_W + v$ where v satisfies the two above hypothesis. Precisely, they say V to be invertible if there is a U such that $U\mu << \mu$ with the further property that $\mu - a.s.$

$$U \circ V = V \circ U = I_W$$

And they found sufficient conditions for this to happen. In **[47]**, Üstünel discovered a necessary condition for this to happen (in a general non-Markovian framework), and we will come back on this below.

The point is that the notion of invertibility of adapted shifts on Wiener space by Üstünel and Zakai, is exactly the restriction to probabilities ν equivalent to the Wiener measure of the notion of stochastic invertibility we introduced above. Indeed for any $v := \int_0^{\cdot} \dot{v}_s ds$ adapted such that (5) holds $\mu - a.s.$, and such that (6) holds, we can define a probability ν equivalent to μ by setting $\mu - a.s.$

$$\frac{d\nu}{d\mu} = \exp\left(-\int_0^1 \dot{v}_s dW_s - \frac{1}{2}\int_0^1 |\dot{v}_s|^2_{\mathbb{R}^d} ds\right)$$

Then, the Girsanov theorem implies that $V := I_W + v$ is the Brownian transform of ν. Conversely for any $\nu \sim \mu$ the associated Brownian transform comes with the shape $V := I_W + v$ where v satisfies the hypothesis of **[53]**. Hence, when restricted to probabilities equivalent to μ the stochastic invertibility (i.e. the invertibility of the Brownian transforms of laws of stochastic diffenrential equations) is equivalent to their notion of invertibility.

Still within the framework of **[53]** hypothesis Üstünel proved the following result (**[47]**) [2] . Let $U := I_W + u$ where $u := \int_0^{\cdot} \dot{u}_s ds$ satisfies (6) and

$$E_\mu\left[\int_0^1 |\dot{u}_s|^2_{\mathbb{R}^d} ds\right] < \infty$$

Further assume that there is an adapted $v := \int_0^{\cdot} \dot{v}_s ds$ which satisfies (6), $\mu - a.s.$ (5), and with the further property that U and V are related by

$$\frac{dU\mu}{d\mu} = \exp\left(-\int_0^1 \dot{v}_s dW_s - \frac{1}{2}\int_0^1 |\dot{v}_s|^2_{\mathbb{R}^d} ds\right) \tag{7}$$

Then, if we note $H(U\mu|\mu)$ the relative entropy which is defined by

$$H(U\mu|\mu) = E_\mu\left[\frac{dU\mu}{d\mu}\ln\frac{dU\mu}{d\mu}\right]$$

[2] In the Markovian case, at least a part of this result already appeared earlier in papers related to stochastic mechanics. Hence, at least part of it may be known at least since the 90's (See Theorem 1.2. of **[34]**, and references therein). However in these earlier works, it seams that the shifts are Markovian, and that this notion is related to the existence of solution, and not to the existence of a unique strong solution which characterizes the invertibility on Wiener space. We also recently learned that this result was related with the main result of **[3]** : the connexion with this formula will not be investigated in this work. We hope that our results can enlighten some results obtained by means of this formula.

we have

$$2H(U\mu|\mu) \leq E_\mu[|u|_H^2]$$

with equality if and only if V is invertible with inverse U. Note that the fact that V is the Brownian transform of $U\mu$ is exactly the hypothesis (7). The notion of invertibility of adapted shifts on Wiener space of **[53]** and the associated criterion of inveribility based on the entropy of **[47]** were the starting points of this thesis. The main goal was to provide some applications of this formula, and to extend it. Hence, in the Part I of this work we provide some general results related to the invertibility and we investigate various applications of the invertibility of adapted shifts on Wiener space, under the hypothesis of **[53]** and **[47]**. In the Part II, we extend the notion of invertibility of adapted shifts on Wiener space of **[53]** to the notion of stochastic invertibility we presented above. We first extend it to probabilities absolutely continuous with respect to the Wiener measures, and we explain how the results of Part I can be extended to probabilities absolutely continuous with respect to the Wiener measure. Then we extend it to stochastic differential equations with dispersion. In Part III we show how the notion of invertibility can fit in stochastic differential geometry by taking the example of the path space with values in a finite dimensional Lie group. Three papers are associated to this work (**[27]**, **[28]**, **[29]**).

Summary and main results

Part I.

Chapter I. We fix the main notations for the whole Part I : we work on a generalization of the classical Wiener space, and we study the notion of invertibility under the hypothesis of **[53]**. We also remark that thanks to a change of measure, the variance of the Girsanov shift associated with a probability equivalent to the Wiener measure may be seen as an entropy (Proposition I.1). This remark will play a key role in the Chapter V.

Chapter II. In this chapter we generalize the criterion of invertibility based on the entropy we recalled above, and we give a very short proof of it (Théorème II.1). We also provide some results of convergence, and in the most general case, we provide an explicit formula for the inverse of a stopped invertible shift (Proposition II.2).

Chapter III. In this section we show that the invertibility is a local property in the usual sense of stochastic calculus. We first prove it under a finite entropy condition thanks to the criterion of invertibility based on the entropy (Theorem III.1). Then we show that this result holds in the most general case (Theorem III.2).

Chapter IV. We prove that pathwise uniqueness holds for the stochastic picture of euclidean quantum mechanics with a great generality. We only investigated the free case, but our method easily extends to the case of potentials. In the finite dimensional case, we prove that the result holds under a finite entropy condition (Theorem IV.2) : the proof relies on the localization of well known results. In the infinite dimensional case (Theorem IV.3), we use a fixed point method wich involves Malliavin calculus, and a localization procedure. One aspect of these results is to provide a consistency to the stochastic picture of euclidean quantum mechanics : indeed the pathwise uniqueness is the stochastic counter part to the determinism of the underlying physic. Moreover these results provides a variational formula for the relative entropy (Proposition IV.3). Then we get a similar formula for an analogous problem of optimal transport (Proposition IV.4).

Chapter V. The main result of this chapter is the Theorem V.3, which extends Shannon's inequality to any abstract Wiener space. We show that Shannon's inequality may be seen as a consequence of information loss in a Gaussian channels.

Part II.

Chapter VI. In this chapter we extend the notion of invertibility to probabilities absolutely continuous but not necessarily equivalent to the Wiener measure. We do it in a very general framework where the temporal structure is poorer than on the classical Wiener space, but where it suffices to get the essential results of stochastic calculus. These structures have been introduced in [**49**] and have been further studied in [**50**]. To be precise we work on an abstract Wiener space where a temporal structure is given by a continuously increasing sequence of projections on the Cameron-Martin space. This poorer structures don't allow some operations, and force to get more fundamentals proofs as the usual ones. In this general framework we prove the existence of the Brownian transform for probabilities absolutely continuous with respect to the Wiener measure (Theorem VI.3). This theorem was proved for measures equivalent to the Wiener measure in [**50**], and part of it also appeared in the same reference, but with a wrong proof. Then we introduce the notion of ν invertibility of the Brownian transform of such a probability ν and we generalize several results within this context. For instance the criterion of invertibility based on the entropy. We also investigate some applications. We show that the innovation conjecture of filtering only depends on the law of the received signal : the innovation conjecture holds if and only if the Brownian transform of its law is stochastically invertible (Theorem VI.9), thus extending a result of [**46**]. We also give a new connexion between invertibility and optimal transport (Theorem VI.8).

Chapter VII. We consider stochastic differential equations of the form

$$dX_t = \alpha_t(X)dW_t + \beta_t(X)dt; X_0 = 0 \tag{8}$$

In this chapter, under a non-degeneracy assumption under which (8) has very similar properties as in the Brownian case, we define the Brownian transform for a law of (1.67), and we investigate several properties related to its invertibility.

Part III.

Chapter VIII. We use the stochastic exponential to lift the Wiener measure to the space of the paths with values in finite dimensional Lie groups. Then we translate the notion of invertibility of adapted shifts within this framework. We show how this notion of invertibility is still related to the existence of a unique strong solution to a stochastic differential equation on the Lie group, and we extend several results within this context. Finally we extend the criterion of invertibility based on the entropy to this space (Theorem VIII.5), which is the main result of the chapter.

Part 1

Invertibility of adapted shifts associated with probabilities equivalent to the Wiener measure

I

Preliminaries and notations

ABSTRACT. **Main notations. Malliavin Calculus. Girsanov shift. Föllmer's theorem, Sobolev's inequality. Talagrand's inequality**

1. Notations

Let (S, H_S, i_S) be an abstract Wiener space (see **[25]**) where S is a separable Banach space, H_S the associated Cameron-Martin space, and i_S the injection of H_S into S which is dense and continuous. In this paper S will be the space where the processes take their values. Indeed we will work on the space $W := C_0([0,1], S)$ of the continuous paths vanishing at 0 with values in S. We recall that W is also a separable Banach space with an associated Cameron-Martin space H defined by

$$H = \left\{ \eta : [0,1] \to H_S, \eta_t = \int_0^t \dot{\eta}_s ds, \int_0^1 |\dot{\eta}_s|_{H_S}^2 ds < \infty \right\}$$

Moreover the scalar product on H is given by

$$< h, k >_H = \int_0^1 < \dot{h}_s, \dot{k}_s >_{H_S} ds$$

for any $h, k \in H$. The injection i of H into W is also dense and continuous and (W, H, i) is also an abstract Wiener space. Let μ be the Wiener measure on $(W, \mathcal{B}(W))$ and let $\mathcal{F}$ be the completion of the Borelian sigma$-$field $\mathcal{B}(W)$ with respect to μ. We still note μ the unique extension of the Wiener measure on $(W, \mathcal{F})$ and we still call it the Wiener measure. We also note $I_W : \omega \in W \to \omega \in W$ the identity map on W. In the sequel I_W will be seen as an equivalence class of $M_\mu((W,\mathcal{F}),(W,\mathcal{F}))$, where $M_\mu((W,\mathcal{F}),(W,\mathcal{F}))$ is the set of the $\mu-$ equivalence classes of mappings from W into itself, which are $\mathcal{F}/\mathcal{F}-$measurable. To cope with adapted processes we need to introduce not only the filtration $\left(\mathcal{F}_t^0\right)$ generated by the coordinate process $t \to W_t$, but also the filtration $(\mathcal{F}_t)$ which is the usual augmentation (see **[6]**) of $\left(\mathcal{F}_t^0\right)$ with respect to the Wiener measure μ. We note $L^0(\mu, H)$ (resp. for a probability law ν equivalent to μ, $L^2(\nu, H)$) the set of the equivalence classes with respect to μ of the measurable $H-$valued mappings $u : W \to H$ (resp. the subset of the $u \in L^0(\mu, H)$ such that $E_\nu\left[|u|_H^2\right] < \infty$). We also set $L_a^0(\mu, H)$ (resp. $L_a^2(\nu, H)$) the subset of the $u \in L^0(\mu, H)$ (resp. of the $u \in L^2(\nu, H)$) such that $t \to \dot{u}_t$ is adapted to $(\mathcal{F}_t)$. Let ν be a probability equivalent to μ, and $t \to B_t$ be a $(\mathcal{F}_t)-$Wiener process on $(W, \mathcal{F}, \nu)$. The abstract stochastic integral (see **[50]**) of a $a \in L_a^0(\mu, H)$ with respect to B will be noted $\delta^B a$. In this context ($W = C_0([0,1], S)$) it can also be written

$$\delta^B a = \int_0^1 \dot{a}_s dB_s$$

For a shift $U := B + u$ where $u \in L_a^0(\mu, H)$ we set

$$\delta^U a := \delta^B a + < u, a >_H$$

In particular for a $u \in L_a^0(\mu, H)$, $\delta^W u = \int_0^1 \dot{u}_t dW_t$. We recall that in the case where W is the classical Wiener space, the abstract stochastic integral is nothing but the usual stochastic integral. For convenience of notations, for any optional time τ with respect to $(\mathcal{F}_t)$ we note $(\pi_\tau a)_. = \int_0^. 1_{[0,\tau]}(s)\,\dot{a}_s ds$ and $a^\tau := \pi_\tau a$. In particular for any $t \in [0,1]$ $(\pi_t a)_. = \int_0^. 1_{[0,t]}(s)\,\dot{a}_s ds$. For a $U \in M_\mu((W,\mathcal{F}),(W,\mathcal{F}))$ and a probability ν equivalent to μ, the image measure of ν by U will be denoted by $U\nu$. Moreover for any random variable L on $(W,\mathcal{F})$ such that $E_\mu[L] = 1$ and $\mu - a.s.$ $L \geq 0$, we will note $L.\mu$ the probability on $(W,\mathcal{F})$ whose density with respect to the Wiener measure is L. To be consistent with **[47]** we set

$$\rho\left(-\delta^W u\right) := \exp\left(-\delta^W u - \frac{|u|_H^2}{2}\right) \tag{1.1}$$

For the sections below Section III we need to consider the Wiener measure on S : we note it $\widehat{\mu}$ and we recall that we then have $W_1\mu = \widehat{\mu}$. We will also consider some integrals over Borelian measures of $S-$valued or H_S valued elements : all these integrals are Bochner integrals. Finally in the whole paper we adopt the convention $\inf(\emptyset) = 1$. We now give the notations of Malliavin calculus and then we give a brief reminder of it (see **[30]**, **[44]**, **[45]**, **[40]**, **[22]**, or **[40]** for much details). We note $\{\delta h, h \in H\}$ (resp.$\{< ., h >, h \in H_S\}$) the isonormal Gaussian field which is the closure of $W^\star$ (resp. of $S^\star$) in $L^p(\mu)$ (resp. in and $L^p(\widehat{\mu})$) and ∇ (resp. D) will denote the Malliavin derivative on W (resp. on S). We also note $D_s\phi$ the density of $\nabla\phi$ i.e. $\nabla_h\phi = \int_0^1 < D_s\phi, \dot{h}_s >_{H_S} ds$ for any $h \in H$. We now recall the construction of the derivative on W, but the construction on S is exactly the same (we don't use the time structure at this point). Let $(k_i)_{i\in\mathbb{N}} \subset H$ be an orthonormal basis of H, let E be a separable Hilbert space and let $(e_i)_{i\in\mathbb{N}} \subset E$ be an orthonormal basis of E. For every $F \in \cap_{p>1} L^p(\mu, E)$ we say that F is a cylindrical function and we note $F \in \mathcal{S}_\mu(E) \subset \cap_{p>1} L^p(\mu, E)$ if there exist a $n \in \mathbb{N}$, $(l_1, ..., l_n) \in (\mathbb{N}^*)^n$, $(k_{l_1}, ..., k_{l_n}) \subset (k_i)_{i\in\mathbb{N}}$ and a f in the Schwartz space of the smooth rapidly decreasing functions $\mathcal{S}(\mathbb{R}^n)$ such that $\mu - as$

$$F = \sum_{i=1}^m f^i(\delta k_{l_1}, ..., \delta k_{l_n})\, e_i$$

If we set

$$\nabla_h F = \frac{d}{d\lambda} F \circ \tau_{\lambda h}\Big|_{\lambda=0}$$

where for any $h \in H$

$$\tau_h : \omega \in W \to \tau_h(\omega) := \omega + h \in W$$

we then have

$$\nabla_h F = \sum_{i=1}^m \sum_{j=1}^n \partial_j f^i(\delta k_{l_1}, ..., \delta k_{l_n}) < h, k_{l_j} >_H e_i$$

By construction, up to a negligible set, for every $\omega \in W$ the mapping defined by $(\nabla F)(\omega) : h \in H \to (\nabla_h F)(\omega) \in E$ is linear and continuous and even Hilbert-Schmidt with the property that $\nabla_h F(\omega) = (\nabla F)(\omega)(h)$. Therefore, by using Hilbert-Schmidt tensor products we have the explicit formula :

$$\nabla F = \sum_{i=1}^m \sum_{j=1}^n \partial_j f^i(\delta k_{l_1}, ..., \delta k_{l_n})\, k_{l_j} \otimes e_i$$

and we have defined a linear operator $\nabla : \mathcal{S}_\mu(E) \subset L^p(\mu, E) \to L^p(\mu, H \otimes E)$. Thanks to the Cameron-Martin theorem it is easy to see that although ∇ is not a closed operator, it is however closable. We still denote by $\nabla : Dom_p(\nabla, E) \subset L^p(\mu, E) \to L^p(\mu, H \otimes E)$ the closure of $\nabla : \mathcal{S}_\mu(E) \subset L^p(\mu, E) \to L^p(\mu, H \otimes E)$ which can be built explicitly in the following way. Let $Dom_p(\nabla, E)$ be the set of the $F \in L^p(\mu, E)$ for which there is a sequence of cylindrical random variables $(F_n)_{n\in\mathbb{N}} \subset \mathcal{S}_\mu(E)$ with the property that $\lim_{n\to\infty} F_n = F$ in $L^p(\mu, E)$ and ∇F_n is Cauchy in $L^p(\mu, H \otimes E)$. Then for any $F \in Dom_p(\nabla, E)$ we can define $\nabla F = \lim_{n\to\infty} \nabla F_n$ which is unique since ∇ is closable. By construction $Dom_p(\nabla, E)$ is the completion of $\mathcal{S}_\mu(E)$ with respect to the norm of the graph associated with ∇ which is defined by $||F||_{p,1;E} = ||F||_{L^p(\mu,E)} + ||\nabla F||_{L^p(\mu,H\otimes E)}$. We note $\mathbb{D}_{p,1}(E)$ the Banach space $Dom_p(\nabla, E)$ endowed with the norm $||F||_{p,1;E}$. Of course ∇ is nothing but the infinite dimensional version of the Sobolev derivative with respect to the Gaussian measure, and $\mathbb{D}_{p,1}(E)$ is the Sobolev space associated with the weak Gross-Sobolev derivative ∇. We define the higher order derivatives and the associated Sobolev spaces by iterating the same procedure. Thus, if $\nabla^{k-1}F \in \mathbb{D}_{p,1}\left(E \otimes H^{\otimes(k-1)}\right)$ we can define $\nabla^k F := \nabla\left(\nabla^{k-1}F\right)$ and the associated Sobolev space $\mathbb{D}_{p,k}(E)$ as being the set of such F equipped with the norm $||F||_{p,k;E} = \sum_{i=0}^{k} ||\nabla^i F||_{L^p\left(\mu, E\otimes H^{\otimes i}\right)}$. In the sequel we will often deal with the case where $E = \mathbb{R}$. Note that in that case, because of the Riesz representation theorem, $H \otimes \mathbb{R} \simeq H$ so that we can identify (with fixed ω) $\nabla F(\omega)$ with a vector of H and we will write $\nabla_h F =< h, \nabla F >_H$. Still in that case we note $\mathbb{D}_{p,1}$ instead of $\mathbb{D}_{p,1}(\mathbb{R})$. Finally we define the so called divergence operator. By the monotone class theorem and from the martingale convergence theorem it is easy to see that $\mathcal{S}_\mu(E)$ is dense in every $L^p(\mu, E)$, $p \geq 1$. Since $\mathcal{S}_\mu(E) \subset \mathbb{D}_{p,1}(E)$, the operator $\nabla : Dom_p(\nabla, E) \subset L^p(\mu, E) \to L^p(\mu, H \otimes E)$ has a dense support. Therefore there is an operator δ, the so called divergence, which is the adjoint of ∇. The domain $Dom_p(\delta, E)$ is defined classically as being the set of the random variables $\xi \in L^p(\mu, H \otimes E)$ such that for any $\phi \in \mathbb{D}_{q,1}(E)$ (where $\frac{1}{p} + \frac{1}{q} = 1$) $E_\mu\left[< \nabla\phi, \xi >_{H\otimes E}\right] \leq c_{p,q}\left(|\phi|_{L^q(\mu,E)}\right)$. For any $\xi \in Dom_p(\delta, E)$, $\delta\xi$ is characterized by the relation $E_\mu\left[< \phi, \delta\xi >_E\right] = E_\mu\left[< \nabla\phi, \xi >_{H\otimes E}\right]$ which holds for any $\phi \in Dom_q(\nabla, E)$. Of course this relation is the infinite dimensional counterpart of the integration by part with respect to the Gaussian measure. Note that the set of the constant H valued random variables is a subset of all the $Dom_p(\delta) := Dom_p(\delta, \mathbb{R})$, and that the Cameron-Martin theorem implies $E_\mu[\phi\delta h] = E\left[< \nabla\phi, h >_H\right]$ for any $h \in H$. Hence it is clear that one may think to this operator as an extension of $\delta : H \to L^p(\mu, H)$, which justifies the notations. The divergence of a u on S can be defined likewise and we note it $< u, . >$.

2. The Girsanov shift

Let ν be a probability which is equivalent to the Wiener measure μ. Then it is well known (see for instance Section 2.6. of **[50]** or **[27]**) that there is a unique $v \in L_a^0(\mu, H)$ such that $t \to W_t + v_t$ is a $(\mathcal{F}_t)$ −abstract Wiener process on $(W, \mathcal{F}, \nu)$ and $\mu - a.s.$

$$\frac{d\nu}{d\mu} = \rho\left(-\delta^W v\right)$$

We say that v (resp. $V := I_W + v$) is the Girsanov drift (resp. shift) associated with ν. We also recall that when $\mu - a.s.$ $\frac{d\nu}{d\mu} = 1 + \delta\alpha$ for a $\alpha \in L_a^0(\mu, H)$ we have $dt \times d\mu - a.s.$ $\dot{v}_t = -\frac{\dot{\alpha}_t}{E_\mu\left[\frac{d\nu}{d\mu}\middle|\mathcal{F}_t\right]}$. If we further assume that $\frac{d\nu}{d\mu} \in \mathbb{D}_{2,1}$ an easy application of the Clark-Ocone formula (see**[45]**) yields

$$\dot{v}_s = -E_\nu\left[D_s \ln \frac{d\nu}{d\mu}\middle|\mathcal{F}_s\right]$$

An old result of Föllmer (see**[14]** and **[15]**) relates the integrability of the Girsanov drift to the relative entropy. In the case of probabilities equivalent to μ, a generalization of this result is the Proposition I.1 (to recover Föllmer result take $\widetilde{\nu} = \mu$ in (2.2) below) which shows that the relative entropy of two probabilities equivalent to the Wiener measure is related to a distance between their Girsanov drifts. In particular it shows that the variance of the Girsanov drift may also be seen as an entropy. We recall that the relative entropy of a probability ν absolutely continuous with respect to a probability $\widehat{\nu}$ is defined by

$$H(\nu|\widehat{\nu}) = E_\nu\left[\ln \frac{d\nu}{d\widehat{\nu}}\right]$$

Proposition I.1. *Let ν and $\widetilde{\nu}$ be two probabilities equivalent to the Wiener measure μ. Then we have*

$$2H(\nu|\widetilde{\nu}) = E_\nu\left[|v - \widetilde{v}|_H^2\right] \tag{2.2}$$

where v (resp. $\widetilde{v}$) is the Girsanov drift associated with ν (resp. with $\widetilde{\nu}$). In particular for any measure ν equivalent to μ with a Girsanov drift v, if we set

$$\frac{d\mu_\nu}{d\mu} := \rho\left(-\delta^W E_\nu[v]\right) \tag{2.3}$$

(see (1.1) for the definition of the right-hand term) we then have

$$2H\left(\nu|\mu_\nu\right) = \mathcal{V}ar_\nu\left(v\right)$$

where $\mathcal{V}ar_\nu\left(v\right) := E_\nu\left[\left|v - E_\nu\left[v\right]\right|_H^2\right]$

Proof: By definition we have

$$\begin{aligned}
E_\nu\left[\ln\frac{d\nu}{d\widetilde{\nu}}\right] &= E_\nu\left[\ln\frac{d\nu}{d\mu}\right] - E_\nu\left[\ln\frac{d\widetilde{\nu}}{d\mu}\right] \\
&= E_\nu\left[-\delta^W v - \frac{|v|_H^2}{2}\right] - E_\nu\left[-\delta^W \widetilde{v} - \frac{|\widetilde{v}|_H^2}{2}\right] \\
&= E_\nu\left[-\delta^V v + \frac{|v|_H^2}{2}\right] - E_\nu\left[-\delta^V \widetilde{v} + < v, \widetilde{v} >_H - \frac{|\widetilde{v}|_H^2}{2}\right] \\
&= E_\nu\left[-\delta^V\left(v - \widetilde{v}\right) + \frac{|v - \widetilde{v}|_H^2}{2}\right]
\end{aligned}$$

If $v - \widetilde{v} \in L_a^2\left(\nu, H\right)$ we have $E_\nu\left[-\delta^V\left(v - \widetilde{v}\right)\right] = 0$ so that equality (2.2) holds. Conversely we assume that $H\left(\nu|\widetilde{\nu}\right) < \infty$ and for each $n \in \mathbb{N}$ we set $\tau_n = \inf\left(\{t \in [0,1] : |\pi_t\left(v - \widetilde{v}\right)|_H > n\}\right)$ with the convention $\inf\left(\emptyset\right) = 1$. Since $v - \widetilde{v} \in L_a^0\left(\mu, H\right)$ we have $\mu - a.s.\ \tau_n \uparrow 1$. Therefore the monotone convergence theorem implies

$$\begin{aligned}
E_\nu\left[|v - \widetilde{v}|_H^2\right] &= E_\nu\left[\lim_{n\to\infty}|\pi_{\tau_n}\left(v - \widetilde{v}\right)|_H^2\right] \\
&= \lim_{n\to\infty} E_\nu\left[|\pi_{\tau_n}\left(v - \widetilde{v}\right)|_H^2\right] \\
&= \lim_{n\to\infty} H\left(\nu|_{\mathcal{F}_{\tau_n}}|\widetilde{\nu}|_{\mathcal{F}_{\tau_n}}\right) \\
&\leq H\left(\nu|\widetilde{\nu}\right)
\end{aligned}$$

Where $\widetilde{\nu}|_{\mathcal{F}_{\tau_n}}$ (resp. $\nu|_{\mathcal{F}_{\tau_n}}$) is the measure induced by $\widetilde{\nu}$ (resp. by ν) on $\mathcal{F}_{\tau_n}$. Hence we proved that $v - \widetilde{v} \in L_a^2\left(\nu, H\right)$ if and only if $H\left(\nu|\widetilde{\nu}\right) < \infty$, and that we always have (2.2). □

Remark I.1. *Consider the function* $g^\lambda\left(\eta\right) := H\left(\eta|\lambda\right)$ *where* η *is any Borelian probability on* $\mathbb{R}^d$ *which is equivalent to Lebesgue measure* λ *on* $\mathbb{R}^d$. *For any* $h \in \mathbb{R}^d$ *the translation* $T_h : x \in \mathbb{R}^d \to x + h \in \mathbb{R}^d$ *is invertible and* λ *is invariant under the action of* T_h *(i.e.* $T_h\lambda = \lambda$*). Thus for any such* η *and any* $h \in \mathbb{R}^d$ $g^\lambda\left(T_h\eta\right) = H\left(T_h\eta|\lambda\right) = H\left(T_h\eta|T_h\lambda\right) = H\left(\eta|\lambda\right) = g^\lambda\left(\eta\right)$. *On the path space the Lebesgue measure is no more defined and one often consider* $g^\mu\left(\nu\right) := H\left(\nu|\mu\right)$ *where* ν *is a Borelian measure. For any* $h \in H$ *let* $\tau_h : \omega \in W \to \omega + h \in W$. *Since* $\tau_h\mu \neq \mu$ *for any* $h \neq 0$, *we generally don't have* $g^\mu\left(\nu\right) := g^\mu\left(\tau_h\nu\right)$. *However the function* $f\left(\nu\right) := H\left(\nu|\mu_\nu\right)$ *has the nice property to be invariant under translations along* H *(just as* g^λ *was on* $\mathbb{R}^d$*). Indeed an easy application of the Cameron- Martin theorem shows that for any* $h \in H$

$$\frac{d\mu_{\tau_h\nu}}{d\mu} = \rho\left(-\delta^W\left(\tau_{-h}\left(E_\nu\left[v\right]\right)\right)\right) \tag{2.4}$$

where v *is the Girsanov drift associated with* ν *and that*

$$\tau_h\mu_\nu = \mu_{\tau_h\nu}$$

Hence for any $h \in H$ *we have*

$$\begin{aligned}
f\left(\tau_h\nu\right) &= H\left(\tau_h\nu|\mu_{\tau_h\nu}\right) \\
&= H\left(\tau_h\nu|\tau_h\mu_\nu\right) \\
&= H\left(\nu|\mu_\nu\right) \\
&= f\left(\nu\right)
\end{aligned}$$

where the last equalities hold since τ_h *is invertible.*

Two straightforward consequences of Proposition I.1 which we will use in the sequel are the following path space version of two inequalities : The Talagrand inequality (see **[42]**, **[45]**, and references therein) and the Sobolev inequality (see **[45]** and **[19]**). The following proofs are well known (see **[45]** and **[47]**). We only give it here for the sake of completeness. However, before we do this we have to recall the following definition :

Definition I.1. *Let ρ and ν be two probabilities on a Wiener space $\widetilde{W}$ (in the sequel we shall consider $\widetilde{W} = S$ or $\widetilde{W} = W$). We then note $\Sigma(\rho,\nu)$ the set of the probabilities on $\left(\widetilde{W}\times\widetilde{W}, \mathcal{B}\left(\widetilde{W}\times\widetilde{W}\right)\right)$ whose first (resp. second) marginal is ρ (resp. ν). A probability $\gamma \in \Sigma(\rho,\nu)$ is said to be the solution of the Monge-Kantorovitch problem if*

$$J(\gamma) := \int_{W\times W} |x-y|_H^2 d\gamma(x,y) = \inf\left(\left\{\int_{W\times W} |x-y|_H^2 d\beta(x,y) : \beta\in\Sigma(\rho,\nu)\right\}\right)$$

The Wasserstein distance between ρ and ν is defined by $\sqrt{d(\rho,\nu)}$, where $d(\rho,\nu) := J(\gamma)$

Proposition I.2. *For any probability ν equivalent to μ we have*

$$d(\nu,\mu) \leq 2H(\nu|\mu) \tag{2.5}$$

Proof: Let V be the Girsanov shift associated with ν. It is a $(\mathcal{F}_t)-$Wiener process on $(W,\mathcal{F},\nu)$. Therefore $(V\times I_W)\,\nu\otimes\nu \in \Sigma(\mu,\nu)$. Together with the definition of $d(\nu,\mu)$ it yields : $d(\nu,\mu) \leq E_\nu\left[|V-I_W|_H^2\right]$. From Proposition I.1 it implies 2.5 □

Proposition I.3. *Let ν be a probability equivalent to μ , and further assume that $\frac{d\nu}{d\mu}\in\mathbb{D}_{2,1}$. Then we have*

$$H(\nu|\mu) \leq J(\nu|\mu)$$

where

$$J(\nu|\mu) := E_\nu\left[\left|\nabla \ln\frac{d\nu}{d\mu}\right|_H^2\right]$$

Proof: Let v be the Girsanov drift associated with ν. Together with Jensen's inequality, Proposition I.1 yields

$$\begin{aligned} 2H(\nu|\mu) &= \int_0^1 E_\nu\left[|\dot{v}_s|_{H_S}^2\right]ds \\ &= \int_0^1 E_\nu\left[\left|E_\nu\left[D_s\ln\frac{d\nu}{d\mu}\middle|\mathcal{F}_s\right]\right|_{H_S}^2\right]ds \\ &\leq \int_0^1 E_\nu\left[\left|D_s\ln\frac{d\nu}{d\mu}\right|_{H_S}^2\right]ds \\ &= E_\nu\left[\left|\nabla\ln\frac{d\nu}{d\mu}\right|_H^2\right] \end{aligned}$$

□

II

Invertibility of adapted shifts on Wiener space

Abstract. **Invertibility of adapted shifts. Definition. From right Invertibility to Invertibility. Explicit formula for the inverse of a stopped shift. Invertibility and Entropy. Üstünel's criterion. A lemma of persistence**

1. Invertibility of adapted shifts

Let $v \in L_a^0(\mu, H)$ be such that $E_\mu\left[\rho\left(-\delta^W v\right)\right] = 1$, and let $\nu := \rho\left(-\delta^W v\right).\mu$ be the probability associated with v. By definition ν is equivalent to μ, and from the Girsanov theorem $V\nu = \mu$. Moreover from Theorem 2.3.1 of **[50]**, for any $u \in L_a^0(\mu, H)$ we have $U\mu << \mu$ where $U := I_W + u$. In particular we have $U\mu << \mu$, $V\nu << \mu$, and $\mu \sim \nu$. It is then routine to check that $U \in M_\mu((W, \mathcal{F}), (W, \mathcal{F}))$, $V \in M_\mu((W, \mathcal{F}), (W, \mathcal{F}))$, and that $V \circ U$ and $U \circ V$ are well defined as elements of $M_\mu((W, \mathcal{F}), (W, \mathcal{F}))$ i.e. the equivalence classes of $U \circ V$ and $V \circ U$ in $M_\mu((W, \mathcal{F}), (W, \mathcal{F}))$ only depend on the equivalence classes of U and V in $M_\mu((W, \mathcal{F}), (W, \mathcal{F}))$. Thus we can define :

Definition II.1. *Let $v \in L_a^0(\mu, H)$ be such that $E_\mu\left[\rho\left(-\delta^W v\right)\right] = 1$. Then $V := I_W + v$ is said to be (globally) invertible with (a global) inverse $U := I_W + u$ where $u \in L_a^0(\mu, H)$, if and only if $\mu - a.s.$*

$$V \circ U = I_W$$

and

$$U \circ V = I_W$$

The next proposition enlighten the hypothesis of **[47]** and is also very useful to get the invertibility from the right invertibility .

Proposition II.1. *Let ν be a probability equivalent to μ, and let $V = I_W + v$ be the Girsanov shift associated with ν. Further assume that there is a $u \in L_a^0(\mu, H)$ such that $U := I_W + u$ is the right inverse of V i.e. $V \circ U = I_W$ $\mu - a.s.$ Then the following assertions are equivalent*

(i) $E_\mu\left[\rho\left(-\delta^W u\right)\right] = 1$
(ii) $U\mu \sim \mu$
(iii) *V is invertible with inverse U (see Definition II.1)*
(iv) $U\mu = \nu$

Proof:

- $(i) => (ii)$: Since $u \in L_a^0(\mu, H)$ and $E_\mu\left[\rho\left(-\delta^W u\right)\right] = 1$ the Girsanov theorem implies $U\left(\rho\left(-\delta^W u\right).\mu\right) = \mu$. Furthermore, for the same reasons, the measure $\rho\left(-\delta^W u\right).\mu$ is a probability equivalent to μ. Hence $\mu = U\left(\rho\left(-\delta^W u\right).\mu\right) \sim U\mu$ i.e. $U\mu \sim \mu$.

- $(ii) => (iii)$ Assume that (ii) holds. We have

$$\begin{aligned} U\mu\left(\{\omega|U\circ V=I_W\}\right) &= \mu\left(U^{-1}\left(\{\omega|U\circ V=I_W\}\right)\right)\\ &= \mu\left(\{\omega|U\circ V\circ U=U\}\right)\\ &\geq \mu\left(\{\omega|V\circ U=I_W\}\right)\\ &= 1\end{aligned}$$

where we have used that $\{\omega|V\circ U=I_W\}\subset\{\omega|U\circ V\circ U=U\}$. Since $U\mu\sim\mu$ we get $\mu-a.s.$ $U\circ V=I_W$.
- $(iii) => (iv)$: Suppose that we have (iii), then we also have $U\circ V\nu=\nu$. Since $V\nu=\mu$ we get $U\mu=\nu$.
- $(iv) => (i)$: Assume that $U\mu=\nu$. Since both v and u are elements of $L_a^0(\mu,H)$, we have $\mu-a.s.$

$$\left(\delta^W v\right)\circ U=\delta^W(v\circ U)+<u,v\circ U>_H$$

so that $\mu-a.s.$

$$\rho\left(-\delta^W v\right)\circ U\rho\left(-\delta^W u\right)=\rho\left(-\delta^W\left(V\circ U-I_W\right)\right)$$

Together with the hypothesis $\mu-a.s.$ $V\circ U=I_W$, it yields $\mu-a.s.$

$$\rho\left(-\delta^W v\right)\circ U\rho\left(-\delta^W u\right)=1$$

Hence we get

$$\begin{aligned} E_\mu\left[\rho\left(-\delta^W u\right)\right] &= E_\mu\left[\frac{1}{\rho\left(-\delta^W v\right)\circ U}\right]\\ &= E_\nu\left[\frac{1}{\rho\left(-\delta^W v\right)}\right]\\ &= 1\end{aligned}$$

□

The Proposition II.2 is an improvement of the Theorem 3.1 of **[48]**. Contrary to this latter, Proposition II.2 provides an explicit formula for the inverse of the stopped shift.

Proposition II.2. *Let ν be a probability which is equivalent to μ with an associated Girsanov shift $V:=I_W+v$, and let τ be a $(\mathcal{F}_t)$ stopping time. We further assume that V is invertible with inverse $U=I_W+u$ (see Definition II.1) where $u\in L_a^0(\mu,H)$. Then $V^\tau:=I_W+\pi_\tau v$ is invertible with inverse*

$$\widetilde{U}:=I_W+\pi_{\tau\circ U}u$$

Proof. Since $(\mathcal{F}_t)$ is right continuous it is well known (see for instance section IV.3 Theorem 59 of **[6]**) that there is a stopping time T with respect to $\left(\mathcal{F}_t^0\right)$ such that $\mu-a.s.$ $T=\tau$. Let $U^\pi:=I_W+\pi_{T\circ U}u$ and $V^T:=I_W+\pi_T v$. Since v is adapted to $(\mathcal{F}_t)$, and from the definition of U^π we have $\mu-a.s.$ $\pi_{T\circ U}(v\circ U)=\pi_{T\circ U}(v\circ U^\pi)$. Since U is the inverse of V (in particular $u=-v\circ U$) this equality can be written

$$\pi_{T\circ U}u=-\pi_{T\circ U}\left(v\circ U^\pi\right)\tag{1.6}$$

Moreover from the Theorem 100.a section IV.4 of **[6]** the fact that for any $\omega\in W$ and for every $s\leq T(U(\omega))$ we have $U_s(\omega)=U_s^\pi(\omega)$, implies that for any $\omega\in W$ $T(U(\omega))=T(U^\pi(\omega))$. Along with (1.6), this implies

$$\pi_{T\circ U}u=-\pi_{T\circ U^\pi}\left(v\circ U^\pi\right)$$

which yields $\mu - a.s.$

$$\begin{aligned} V^T \circ U^\pi - I_W &= (\pi_T v) \circ U^\pi + \pi_{T\circ U} u \\ &= (\pi_T v) \circ U^\pi - \pi_{T\circ U^\pi} (v \circ U^\pi) \\ &= 0 \end{aligned}$$

i.e. U^π is the almost sure right inverse of V^T. We set $u^\pi = \pi_{T\circ U} u$ so that $U^\pi = I_W + u^\pi$. Since U is $(\mathcal{F}_t)$ −adapted, $u^\pi \in L^0_a(\mu, H)$. On the other hand, since U is invertible we know from Proposition II.1 that $E_\mu\left[\rho\left(-\delta^W u\right)\right] = 1$ so that $t \to \rho\left(\delta^W \pi_t u\right)$ is a martingale. Hence we can apply the Doob's optional stopping theorem from which we obtain $\mu - a.s.$

$$E_\mu\left[\rho\left(-\delta^W u^\pi\right)\right] = E_\mu\left[\rho\left(-\delta^W \pi_{T\circ U} u\right)\right] = E_\mu\left[E_\mu\left[\rho\left(-\delta^W u\right)\Big|\mathcal{F}_{\tau\circ U}\right]\right] = 1$$

Therefore Proposition II.1 implies that U^π is the inverse of V^T. Since $\mu - a.s$, $V^\tau = V^T$ and $U^\pi = \widetilde{U}$, we also have that V^τ is invertible with inverse $\widetilde{U}$. □

As a first consequence we have

Lemma II.1. *Let $v \in L^0_a(\mu, H)$ be such that $E_\mu\left[\rho\left(-\delta^W v\right)\right] = 1$, and let $(\tau_n)_{n\in\mathbb{N}}$ be a sequence of optional times such that $\mu - a.s.$ $\tau_n \uparrow 1$. Further assume that there is a sequence $(u^n) \subset L^0_a(\mu, H)$ such that for each $n \in \mathbb{N}$ $V^n := I_W + \pi_{\tau_n} v$ is invertible with inverse $U^n := I_W + u^n$. Then (u^n) converges in $L^0_a(\mu, H)$*

Proof: For convenience of notations we note $v^n := \pi_{\tau_n} v$. By definition and by Proposition II.2, for each $n \geq m$ we have $\mu - a.s.$

$$v^m \circ U^m = v^m \circ U^n$$

This yields

$$\begin{aligned} \mu\left(|u^n - u^m|_H > \epsilon\right) &= \mathrm{E}_\mu\left[1_{|u^n - u^m|_H > \epsilon}\right] \\ &= \mathrm{E}_\mu\left[1_{|v^n \circ U^n - v^m \circ U^m|_H > \epsilon}\right] \\ &= \mathrm{E}_\mu\left[1_{|v^n \circ U^n - v^m \circ U^n|_H > \epsilon}\right] \\ &= \mathrm{E}_\mu\left[1_{|v^n \circ U^n \circ V^n - v^m \circ U^n \circ V^n|_H > \epsilon}\rho\left(-\delta^W v^n\right)\right] \\ &= \mathrm{E}_\mu\left[1_{|v^n - v^m|_H > \epsilon}\rho\left(-\delta^W v^n\right)\right] \end{aligned}$$

On the other hand $\left(\rho\left(-\delta^W v^n\right)\right)_{n\in\mathbb{N}}$ and $(v^n)_{n\in\mathbb{N}}$ converge almost surely to $\rho\left(-\delta^W v\right)$ and v respectively. Since $\left(\rho\left(-\delta^W v^n\right)\right)_{n\in\mathbb{N}}$ is uniformly integrable, the dominated convergence theorem implies that (u^n) converges in $L^0(\mu, H)$ □

Remark II.1. *Of course since the norm $|.|_W$ is weaker than $|.|_H$ the convergence of (u^n) in $L^0_a(\mu, H)$ to a $u \in L^0_a(\mu, H)$ also implies the convergence in probability of (U^n) to $U := I_W + u$.*

2. Invertibility and Entropy

We set

Definition II.2. *Let ν be a probability equivalent to μ, then we set*

$$\mathcal{R}_a(\mu, \nu) = \left\{u \in L^2_a(\mu, H) : U\mu = \nu \text{ where } U := I_W + u\right\}$$

and we say that a $u \in L^2_a(\mu, H)$ represents ν if $u \in \mathcal{R}_a(\mu, \nu)$

The Theorem II.1 extends the Theorem 7 of [**47**]. Our proof is straightforward and contrary to [**47**] we neither use the innovation process, nor the dual predicable projection.

Theorem II.1. *Let ν be a probability such that $\nu \sim \mu$ (i.e. equivalent) and $H(\nu|\mu) < \infty$, and let $V = I_W + v$ be the Girsanov shift associated with ν. Then for every $u \in \mathcal{R}_a(\mu,\nu)$ (see Definition II.2), the following inequality holds*

$$2H(\nu|\mu) \leq E_\mu\left[|u|_H^2\right] \tag{2.7}$$

Moreover, the equality occurs in (2.7) if and only if $V = I_W + v$ is invertible with inverse $U = I_W + u$ (see Definition II.1), and in that case we also have $E_\mu\left[\rho\left(-\delta^W u\right)\right] = 1$

Proof. From Proposition I.1 the finite entropy condition is equivalent to the finite energy condition $E_\nu\left[|v|_H^2\right] < \infty$. Since $V = I_W + v$ it implies

$$E_\nu\left[\delta^V v\right] = 0 \tag{2.8}$$

On the other hand, together with the hypothesis $U\mu = \nu$, the finite energy condition also implies $v \circ U \in L_a^2(\mu, H)$ so that

$$E_\mu\left[\delta^W (v \circ U)\right] = 0 \tag{2.9}$$

Therefore, by putting together (2.8) and (2.9) we obtain

$$\begin{aligned} E_\nu\left[|v|_H^2\right] &= E_\nu\left[\delta^V v\right] - E_\nu\left[\delta^W v\right] \\ &= -E_\mu\left[\delta^U (v \circ U)\right] \\ &= -E_\mu\left[\delta^W (v \circ U)\right] - E_\mu\left[< u, v \circ U >_H\right] \\ &= -E_\mu\left[< u, v \circ U >_H\right] \end{aligned}$$

so that Proposition I.1 yields

$$2H(\nu|\mu) = -E_\mu\left[< u, v \circ U >_H\right] \tag{2.10}$$

By applying the Cauchy-Schwarz inequality, the result follows from (2.10). Indeed we have

$$\begin{aligned} 2H(\nu|\mu) &= -E_\mu\left[< u, v \circ U >_H\right] \\ &\leq \sqrt{E_\mu\left[|u|_H^2\right]}\sqrt{E_\mu\left[|v \circ U|_H^2\right]} \\ &= \sqrt{E_\mu\left[|u|_H^2\right]}\sqrt{E_\nu\left[|v|_H^2\right]} \\ &= \sqrt{E_\mu\left[|u|_H^2\right]}\sqrt{2H(\nu|\mu)} \end{aligned}$$

Moreover form the case of equality in the Cauchy-Schwarz inequality, it is easy to see that the equality occurs in (2.7) if and only if $\mu - a.s.$ $u + v \circ U = 0$ i.e. if and only if $\mu - a.s.$ $V \circ U = I_W$. The result then follows directly form Proposition II.1. □

As a consequence we get

Lemma II.2. *Let ν be a probability such that $\nu \sim \mu$ (i.e. equivalent) and $H(\nu|\mu) < \infty$. Further assume that there are two sequences $(u^n)_{n\in\mathbb{N}} \subset L_a^0(\mu, H)$, and $(v^n)_{n\in\mathbb{N}} \subset L_a^0(\mu, H)$ such that :*

(i) $E_\mu[L_n] = 1$ for every $n \in \mathbb{N}$, where $L_n := \rho\left(-\delta^W v^n\right)$.
(ii) For every $n \in \mathbb{N}$ the shift $V^n := I_W + v^n$ is invertible with inverse $U^n := I_W + u^n$ (see Definition II.1)
(iii) $(L_n \ln L_n)_{n\in\mathbb{N}}$ is bounded in $L^1(\mu)$ or $(u^n)_{n\in\mathbb{N}}$is bounded in $L^2(\mu, H)$.
(iv) $(L_n)_{n\in\mathbb{N}}$ converges weakly in $L^1(\mu)$ to $L := \frac{d\nu}{d\mu}$.
(v) $H(\nu|\mu) = \lim_{n\to\infty} E_\mu[L_n \ln L_n]$
(vi) (u^n) converges in $L^0(\mu, H)$ to a $\widetilde{u}$

Then $\widetilde{u} \in L_a^2(\mu, H)$ and $\widetilde{U} = I_W + \widetilde{u}$ is the almost sure inverse of $V = I_W + v$.

PROOF. First note that (i) means that v^n (resp. V^n) is the Girsanov drift (resp. shift) associated with the probability ν^n defined by $\nu^n := L_n.\mu$. From Proposition I.1 we then have

$$2E_\mu [L_n \ln L_n] = E_\mu \left[L_n |v^n|_H^2\right]$$

Moreover, since U^n is the inverse of V^n we have $u^n = -v^n \circ U^n$, and by Proposition II.1 we also have $U^n \mu = \nu^n = L_n.\mu$. This yields

$$E_\mu \left[|u^n|_H^2\right] = E_\mu \left[|v^n \circ U^n|_H^2\right] = E_\mu \left[L_n |v^n|_H^2\right] = 2E_\mu [L_n \ln L_n] \tag{2.11}$$

which shows that the two hypothesis of (iii) are equivalent. By taking the limit in (2.11), the hypothesis (v) yields

$$\lim_{n\to\infty} E_\mu \left[|u^n|_H^2\right] = 2H(\nu|\mu) \tag{2.12}$$

Let us now assume that (u^n) is bounded in $L^2(\mu, H)$. Since $L^2(\mu, H)$ an Hilbert space it is also reflexive. Therefore we can extract a sequence $(u^n, v^n)_{n\in\mathbb{N}}$ with the property that (u^n) converges in the weak topology of $L^2(\mu, H)$ to a $u \in L^2(\mu, H)$. Let $\pi^{\mathcal{F}} : L^2(\mu, H) \to L^2(\mu, H)$ be the projection of $L^2(\mu, H)$ on the closed subspace $L_a^2(\mu, H)$. Since it is linear and strongly continuous, a classical theorem of functional analysis directly implies that it is also weakly continuous. Hence $\pi^{\mathcal{F}} u = u$ and $u \in L_a^2(\mu, H)$. On the other hand, let $\widetilde{u}$ be the limit of (u^n) in $L^0(\mu, H)$. We now show that $\widetilde{u} = u$. Indeed let $X \in L^\infty(\mu)$ and $h \in L^2(\mu, H)$. Then $Xh \in L^2(\mu, H)$ where $(Xh)(\omega) := X(\omega) h_.(\omega)$. By definition of the weak convergence of (u^n) we then have

$$\begin{aligned}
\lim_{n\to\infty} E_\mu [< u^n, h >_H X] &= \lim_{n\to\infty} E_\mu [< u^n, Xh >_H] \\
&= E_\mu [< u, Xh >_H] \\
&= E_\mu [< u, h >_H X]
\end{aligned}$$

This means that $(< u^n, h >_H)$ converges weakly in $L^1(\mu)$. The Dunford-Pettis criterion (see **[6]**) then implies that for any $h \in L^2(\mu, H)$ the family $(< u^n, h >_H)$ is uniformly integrable. Thus, from the dominated convergence theorem, and by definition of the weak convergence, for any $h \in L^2(\mu, H)$ we have

$$\begin{aligned}
E_\mu [< \widetilde{u}, h >_H] &= E_\mu \left[\lim_{n\to\infty} < u^n, h >_H\right] \\
&= \lim_{n\to\infty} E_\mu [< u^n, h >_H] \\
&= E_\mu [< u, h >_H]
\end{aligned}$$

This shows that $u = \widetilde{u}$, from which it can be seen that u satisfies the hypothesis of Theorem II.1. Indeed, note that together with (iv), the dominated convergence theorem yields

$$\begin{aligned}
E_{U\mu} [\exp(il)] &= E_\mu [\exp(il \circ U)] \\
&= E_\mu \left[\lim_{n\to\infty} \exp(il \circ U^n)\right] \\
&= \lim_{n\to\infty} E_\mu [\exp(il \circ U^n)] \\
&= \lim_{n\to\infty} E_\mu [L_n \exp(il)] \\
&= E_\mu [L \exp(il)] \\
&= E_\nu [\exp(il)]
\end{aligned}$$

for any $l \in W^\star$. Hence we have both $U\mu = \nu$ and $u \in L_a^2(\mu, H)$ so that we are allowed to apply Theorem II.1. From this we obtain

$$2H(\nu|\mu) \le E_\mu \left[|u|_H^2\right] \tag{2.13}$$

On the other hand, since $|.|_{L^2(\mu,H)}$ is both convex and strongly lower semi-continuous, it is also lower semi-continuous in the weak topology of $L^2(\mu, H)$ so that $E_\mu\left[|u|_H^2\right] \leq \lim_{n\to\infty} E_\mu\left[|u^n|_H^2\right]$. Along with equation (2.12), this yields

$$2H(\nu|\mu) \geq E_\mu\left[|u|_H^2\right] \tag{2.14}$$

From (2.13) and (2.14) we obtain

$$2H(\nu|\mu) = E_\mu\left[|u|_H^2\right] \tag{2.15}$$

Since u satisfies the hypothesis of Theorem II.1, equation (2.15) implies that $U = I_W + u$ is the inverse of $V = I_W + v$. □

III

Local properties

Abstract. **Local Invertibility. The problem of equivalence with invertibility. Equivalence under finite energy condition. Equivalence in the general case**

1. Local invertibility : definition and problem

In this chapter we prove that the invertibility of the Girsanov shifts associated with probabilities equivalent to the Wiener measure is a local property in the usual sense of stochastic calculus. We first set the definition of local invertibility :

Definition III.1. *Let $v \in L_a^0(\mu, H)$ be such that $E_\mu\left[\rho\left(-\delta^W v\right)\right] = 1$ and let $V := I_W + v$. V is said to be locally invertible if there is a sequence $(u^n)_{n\in\mathbb{N}} \subset L_a^0(\mu, H)$ and a sequence $(\tau_n)_{n\in\mathbb{N}}$ of $(\mathcal{F}_t)$ −optional times such that $\mu - a.s.$ $\tau_n \uparrow 1$ and for each $n \in \mathbb{N}$, $V^n := I_W + \pi_{\tau_n} v$ is invertible (see Definition II.1) with inverse $U^n := I_W + u^n$.*

First note that any shift which is globally invertible is also locally invertible. Indeed, by taking the sequence $(T_n)_{n\in\mathbb{N}}$ defined by $T_n = 1-1/n$ for any $n \in \mathbb{N}$, the Proposition III.1 appears as a direct consequence of Proposition II.2. However Proposition III.1 is trivial to prove directly, and we only put it here for the sake of completeness.

Proposition III.1. *Let $v \in L_a^0(\mu, H)$ be such that $E_\mu\left[\rho\left(-\delta^W v\right)\right] = 1$, and let $\nu := \rho\left(-\delta^W v\right).\mu$ be the probability law associated with v. Further assume that $V := I_W + v$ is invertible. Then V is locally invertible*

Conversely we will see that the local invertibility implies the invertibility in the general case.

2. Solution under finite energy condition

Theorem III.1. *Let $v \in L_a^0(\mu, H)$ be such that $E_\mu\left[\rho\left(-\delta^W v\right)\right] = 1$, and $\nu := \rho\left(-\delta^W v\right).\mu$. Further assume that v satisfies the finite energy condition (i.e. $v \in L_a^2(\nu, H)$), or equivalently that ν has finite entropy with respect to the Wiener measure (i.e. $H(\nu|\mu) < \infty$). Then V is locally invertible (see Definition III.1) if and only if it is (globally) invertible (see Definition II.1).*

Proof. We already recalled in Proposition III.1 that if V is globally invertible, it is also locally invertible. Hence, we now assume that $V = I_W + v$, which is also the Girsanov shift associated with ν, is locally invertible, and we want to prove that it is globally invertible. Hence we assume that there is a sequence $(\tau_n)_{n\in\mathbb{N}}$ of $(\mathcal{F}_t)$ optional times announcing 1, and a sequence $(u^n) \subset L_a^0(\mu, H)$, such that for every $n \in \mathbb{N}$ $U^n := I_W + u^n$ is the almost sure inverse of V^n where $V^n := I_W + \pi_{\tau_n} v$. For convenience of notations we set $L := \frac{d\nu}{d\mu}$, $L_n := E_\mu\left[L|\mathcal{F}_{\tau_n}\right]$, $\nu^n := L_n.\mu$ and $v^n := \pi_{\tau_n} v$. In particular $L_n = \rho\left(-\delta^W v^n\right)$ and v^n is the Girsanov drift associated with ν^n. We want to apply Lemma II.2 with (u^n) and (v^n). Thus, we have to check that (u^n) and (v^n) satisfy the hypothesis $(i - vi)$ of Lemma II.2. From Lemma II.1 (vi) (u^n) converges in probability to a $\widetilde{u} \in L^0(\mu, H)$ and we have (i). Note that (ii, iii, v) are obvious and that we even have that $L_n \to L$ strongly in $L^1(\mu)$. Hence to prove (vi), we just have to show the uniform integrability of

$(L_n \ln L_n)_{n\in\mathbb{N}}$. Thanks to the Jensen's inequality we have :

$$\begin{aligned} L_n \ln L_n &= E_\mu\left[L|\mathcal{F}_{\tau_n}\right] \ln E_\mu\left[L|\mathcal{F}_{\tau_n}\right] \\ &\leq E_\mu\left[L \ln L|\mathcal{F}_{\tau_n}\right] \end{aligned}$$

Since for any x which is positive $|xlnx| \leq xlnx + 2e^{-1}$, we get $\mu - a.s.$

$$|L_n \ln L_n| \leq E_\mu\left[L \ln L|\mathcal{F}_{\tau_n}\right] + 2e^{-1} \tag{2.16}$$

Together with (2.16), the uniform integrability of $(E_\mu\left[L \ln L|\mathcal{F}_{\tau_n}\right])_{n\in\mathbb{N}}$, implies the uniform integrability of $(L_n \ln L_n)_{n\in\mathbb{N}}$ i.e. (v) is proved. We now check that (u_n) is bounded in $L^2(\mu, H)$. From Theorem II.1, and Jensen's inequality, this is the case. Indeed :

$$\begin{aligned} E_\mu\left[|u^n|_H^2\right] &= 2E_\mu\left[L_n \ln L_n\right] \\ &= 2E_\mu\left[E_\mu\left[L|\mathcal{F}_{\tau_n}\right] \ln E_\mu\left[L|\mathcal{F}_{\tau_n}\right]\right] \\ &\leq 2E_\mu\left[E_\mu\left[L \ln L|\mathcal{F}_{\tau_n}\right]\right] \\ &= 2E_\mu\left[L \ln L\right] \\ &= 2H\left(\nu|\mu\right) \\ &< \infty \end{aligned}$$

The sequences (u^n) and (v^n) satisfy the hypothesis $(i - vi)$ of Lemma II.2, so that V is globally invertible with inverse $\widetilde{U} := I_W + \widetilde{u}$. □

3. General case

The next theorem completely solves the problem of the equivalence between invertibility and local invertibility.

Theorem III.2. *Let $v \in L_a^0(\mu, H)$ be such that $E_\mu\left[\rho\left(-\delta^W v\right)\right] = 1$. Then $V := I_W + v$ is locally invertible if and only if it is invertible.*

Proof: The sufficiency is obvious by taking $(\tau_n := 1 - 1/n)$ and applying Proposition II.2. Conversely we henceforth assume that V is locally invertible. By hypothesis there is a sequence (τ_n) of optional times and a sequence $(u^n) \subset L_a^0(\mu, H)$ such that for each $n \in \mathbb{N}$ $V^n := I_W + \pi_{\tau_n} v$ is invertible with inverse $U^n := I_W + u^n$ and $\mu - a.s.$ $\tau_n \uparrow 1$. By Lemma II.1 (u^n) converges in $L_a^0(\mu, H)$. We note $u \in L_a^0(\mu, H)$ this limit and $U := I_W + u$. We will show that V is invertible with inverse U. For convenience of notations we set $v^n := \pi_{\tau_n} v$, $L_n := \rho\left(-\delta^W v^n\right)$, $L := \rho\left(-\delta^W v\right)$ and $\nu := L.\mu$. From Doob's optional stopping theorem $L_n := E_\mu\left[L|\mathcal{F}_{\tau_n}\right]$ so that (L_n) is uniformly integrable and converges to L in $L^1(\mu)$. On the other hand (see Remark II.1), U^n converges to U in probability. Therefore the dominated convergence theorem yields

$$\begin{aligned} E_\nu\left[e^{il}\right] &= E_\mu\left[Le^{il}\right] \\ &= \lim_{n\to\infty} E_\mu\left[L_n e^{il}\right] \\ &= \lim_{n\to\infty} E_\mu\left[e^{il\circ U^n}\right] \\ &= E_\mu\left[e^{il\circ U}\right] \end{aligned}$$

for any $l \in W^\star$ i.e. $U\mu = \nu$. Thus from Proposition II.1 we know that V is invertible with inverse U if and only if U is the almost sure right inverse of V (i.e. $\mu - a.s.$ $V \circ U = I_W$ or equivalently $\mu - a.s.$ $u + v \circ U = 0$). As we shall see we can show this last result thanks to Lusin's theorem. Let $c > 0$ we have

$$\begin{aligned} \mu\left(|v \circ U + u^n|_W > c\right) &= \mu\left(|v \circ U - v^n \circ U^n|_W > c\right) \\ &\leq \mu\left(|v \circ U - v \circ U^n|_H > \frac{c}{2}\right) + E_\mu\left[L_n 1_{|v \ -v^n|_H > \frac{c}{2}}\right] \end{aligned}$$

Let $\alpha > 0$, the dominated convergence theorem implies the existence of a N_1 such that for any $n > N_1$

$$E_\mu \left[L_n 1_{|v \ -v^n|_H > \frac{c}{2}} \right] < \alpha/2$$

To control $\nu \left(|v \circ U - v \circ U^n|_H > \frac{c}{2} \right)$ we will use Lusin's theorem from which we know the existence of a compact set $K_\alpha \subset W$ such that $\nu(K_\alpha) \geq 1 - \alpha/8$, and v is uniformly continuous on K_α. We set

$$\Omega_n = \left\{ \omega : |v \circ U - v \circ U^n|_H > \frac{c}{2}, (U, U^n) \in K_\alpha \times K_\alpha \right\}$$

We then have

$$\begin{aligned} \mu \left(|v \circ U - v \circ U^n|_H > \frac{c}{2} \right) &\leq \mu(\Omega_n) + \mu(U \notin K_\alpha) + \mu(U^n \notin K_\alpha) \\ &= \mu(\Omega_n) + \nu(\omega \notin K_\alpha) + E_\mu \left[\rho(-\delta v^n) 1_{\omega \notin K_\alpha} \right] \end{aligned}$$

By definition of K_α $\nu(\omega \notin K_\alpha) < \alpha/8$. Moreover, by the dominated convergence theorem

$$\lim_{n \to \infty} E_\mu \left[\rho(-\delta v^n) 1_{\omega \notin K_\alpha} \right] = \nu(\omega \notin K_\alpha) < \frac{\alpha}{8}$$

Thus there is a N_2 such that for any $n > N_2$ $E_\mu \left[\rho(-\delta v^n) 1_{\omega \notin K_\alpha} \right] < \alpha/4$. We then have for any $n > \sup(\{N_1, N_2\})$

$$\mu(|v \circ U + u^n|_W > c) \leq \frac{7\alpha}{8} + \mu(\Omega_n) \tag{3.17}$$

On the other hand, the uniform continuity of v on K_α yields the existence of a $\beta_{\alpha,c}$ such that $|u^n - u|_H < \beta_{\alpha,c}$ and $(U^n, U) \in K_\alpha \times K_\alpha$ imply $|v \circ U - v \circ U^n|_H < \frac{c}{2}$. To control the last term of (3.17) we then set

$$\widetilde{\Omega}_n = \{ \omega \in W : |u^n - u| < \beta_{\alpha,c} \}$$

In particular $\Omega_n \cap \widetilde{\Omega}_n = \emptyset$ so that we obtain

$$\begin{aligned} \mu(\Omega_n) &\leq \mu\left(\Omega_n \cap \widetilde{\Omega}_n \right) + \mu\left(\Omega_n \cap \left(\widetilde{\Omega}_n^c \right) \right) \\ &= \mu\left(\Omega_n \cap \left(\widetilde{\Omega}_n^c \right) \right) \\ &\leq \mu\left(\widetilde{\Omega}_n^c \right) \\ &= \mu(|u_n - u| > \beta_{\alpha,c}) \end{aligned}$$

Since $u_n \to u$ in $L^0_a(\mu, H)$ there is a N_3 such that for any $n > N_3$ $\mu(\Omega_n) < \alpha/8$. Therefore for any $n > \sup(\{N_1, N_2, N_3\})$ we have

$$\mu(|v \circ U + u^n|_W > c) < \alpha$$

this proves that $u^n \to -v \circ U$ in $L^0(\mu, H)$. By unicity of the limit $u + v \circ U = 0$ i.e. U is the right inverse of V. But we already showed that this implies the invertibility of V which is the result. □

We now state the result which will be used in the applications :

Corollary III.1. *Let $v \in L^0_a(\mu, H)$ be such that $E_\mu \left[\rho\left(-\delta^W v \right) \right] = 1$. Assume that there is a sequence $(v^n) \subset L^0_a(\mu, H)$ with the property that for each $n \in \mathbb{N}$*

$$E_\mu \left[\rho\left(-\delta^W v^n \right) \right] = 1$$

and such that $V^n := I_W + v^n$ is invertible. Further assume that there is a sequence of $(\mathcal{F}_t)$ stopping times (τ_n) such that $\mu - a.s.$ $\tau_n \uparrow 1$ and for each $n \in \mathbb{N}$ $\mu - a.s.$

$$\pi_{\tau_n} v = \pi_{\tau_n} v^n$$

Then $V := I_W + v$ is invertible

Proof: By Proposition II.2 for each $n \in \mathbb{N}$ the shift defined by $\widetilde{V}^n := I_W + \pi_{\tau_n} v^n$ is invertible. Therefore V is locally invertible. Hence Theorem III.2 yields the invertibility of V. □

IV

Applications to stochastic mechanics: strong solutions for some Markovian equations

ABSTRACT. **Locally bounded Markovian shifts are invertible. Free Schrödinger shifts. Pathwise uniqueness for quantum euclidean mechanics : finite dimensional case, infinite dimensional case. Invertibility of free Schrödinger shifts and optimal transport.**

1. A sufficient condition for the invertibility of Markovian shifts with states in $\mathbb{R}^d$

In this section we will only consider the case where $S = \mathbb{R}^d$ for a $d \in \mathbb{N}$. The main result of this section is Theorem IV.1 which is a sufficient condition of invertibility for Markovian shifts. From the main result of **[54]** any shift which is both Markovian and bounded is invertible. Here we give a local version of this fact. Note that this extension is different from those of **[24]**.

Definition IV.1. *A $v := \int_0^\cdot \dot{v}_s ds \in L_a^0(\mu, H)$ is said to be locally bounded if there is a sequence of $(\mathcal{F}_t)-$ stopping times (τ_n) such that $\mu - a.s.$, $\tau_n \uparrow 1$ and such that for each $n \in \mathbb{N}$ we have $\mu - a.s.$*

$$\sup_{s \leq \tau_n} |\dot{v}_s|_{H_S} < \infty$$

Proposition IV.1 will enable us to use the notion of Definition IV.1.

Proposition IV.1. *Let $v \in L_a^0(\mu, H)$ and*

$$\sigma_n := \inf\left(\left\{t : \sup_{s\in[0,t]} |\dot{v}_s| > n\right\}\right) \wedge 1$$

Then v is locally bounded (see Definition IV.1) if and only if $\mu - a.s.$ $\sigma_n \uparrow 1$

Proof: If $\sigma_n \uparrow 1$ $\mu - a.s.$ the sequence (σ_n) satisfies the hypothesis of Definition IV.1. Conversely we assume that v is locally bounded, and we define Ω to be the set of the $\omega \in W$ such that $\tau_n(\omega) \uparrow 1$ and such that for all $n \in \mathbb{N}$, $\sup_{s\leq\tau_n(\omega)} |\dot{v}_s(\omega)| < \infty$. From the hypothesis $\mu(\Omega) = 1$. Given $\omega \in \Omega$ and $\epsilon \in [0, 1)$ there is a $n_0 \in \mathbb{N}$ and a $K > 0$ such that $\tau_{n_0}(\omega) > \epsilon$ and

$$\sup_{s\in\left[0,\tau_{n_0}(\omega)\right]} |\dot{v}_s| < K$$

Let $m_0 \in \mathbb{N}$ be such that $m_0 > K$. Then $\sup_{s\in[0,\tau_{n_0}(\omega))} |\dot{v}_s| < K < m_0$ so that $\sigma_{m_0}(\omega) \geq \tau_{n_0}(\omega) > \epsilon$. Since $\sigma_n(\omega)$ increases, this implies that $\sigma_n(\omega) \uparrow 1$. $\square$

Under a mild condition, the next proposition shows that a Markovian shift which is locally bounded in the sense of Definition IV.1 is invertible.

Theorem IV.1. *Let $v \in L_a^0(\mu, H)$ be such that $E_\mu\left[\rho\left(-\delta^W v\right)\right] = 1$. Further assume that there is a measurable $b : [0, 1) \times \mathbb{R}^d \to \mathbb{R}^d$ such that $\mu - a.s.$ for each $t \in [0, 1)$*

$$v_t = \int_0^t b(s, W_s) ds$$

and that v is locally bounded in the sense of Definition IV.1. Then $V := I_W + v$ is invertible.

Proof: Let (σ_n) be as in Proposition IV.1. Since v is locally bounded $\mu - a.s.$ $\sigma_n \uparrow 1$. For each $T \in [0,1)$ and $n \in \mathbb{N}$ we set

$$v^T := \int_0^{\cdot} b(t, W_t) 1_{t \leq T} dt$$

and

$$v^{n,T} := \int_0^{\cdot} b^{n,T}(t, W_t) dt$$

where

$$b^{n,T}(t,x) := b(t,x) 1_{|b(t,x)|<n} 1_{t \leq T}$$

Since $b^{n,T}(t,x)$ is both measurable and bounded the main result of **[54]** yields the existence of a strong solution for the equation

$$dX_t = dB_t - b^{n,T}(t, W_t) dt \tag{1.18}$$

Thus $V^{n,T} := I_W + v^{n,T}$ is right invertible with an inverse $U^{n,T} := I_W + u^{n,T}$(note that since $b^{n,T}$ is bounded the condition $u^{n,T} \in L^0_a(\mu, H)$ is filled). Moreover the fact that $b^{n,T}$ is bounded, together with the Novikov criterion, yields $E_\mu\left[\rho(-\delta^W u^{n,T})\right] = 1$. From Proposition II.1 it yields the invertibility of each $V^{n,T}$. On the other hand from the hypothesis we obviously have $\pi_{\sigma_n} v^T = \pi_{\sigma_n} v^{n,T}$. Hence Corollary III.1 implies the invertibility of $V^T := I_W + v^T$ for each $T < 1$. In particular V is locally invertible (take $\tau_n = 1 - 1/n$) and therefore invertible (Theorem III.2). $\square$

2. Invertibility of free Schrödinger shifts

Let $\widehat{\nu}$ be a probability equivalent to $\widehat{\mu}$, it is well known (see **[16]**) that

$$H(\widehat{\nu}|\widehat{\mu}) = \inf\left(\{H(\nu|\mu) : W_1\nu = \widehat{\nu}\}\right) \tag{2.19}$$

and that the optimum is attained by the probability ν which is defined by

$$\frac{d\nu}{d\mu} = \frac{d\widehat{\nu}}{d\widehat{\mu}} \circ W_1$$

Moreover ν is an h-path process in the sense of Doob (see **[16]** and the references therein) so that the Girsanov drift associated with the optimal ν is Markovian. Such a process may be seen as a particular Schrödinger Bridge (see **[16]** and alternatively **[61]** and references therein). Thus it is connected to stochastic mechanics (see **[37]**, **[38]**, **[39]**) and to stochastic control problems (see **[35]**, **[36]**, **[34]** and references therein). The connection of Schrödinger Bridges with stochastic mechanics was clearly shown in **[61]** and **[62]**, and is trivial to see in the special case of h-path processes. Moreover it is known for a long time that such mechanics are related to stochastic control both through Yasue's approach(**[58]**) and through the Guerra-Morato(-Nelson) approach (see **[18]**, and **[38]**). Consider now the equation

$$dU_t = dW_t - \dot{v}_t \circ U dt; U_0 = 0 \tag{2.20}$$

where v is the Girsanov drift associated with the optimal probability ν. As it appears clearly from **[61]** and **[17]** a solution of (2.20) may been interpreted physically as a free euclidean quantum (time imaginary) particle starting from the origin whose final marginal is empirically estimated by $\widehat{\nu}$. In this context the equivalences investigated in **[17]** show that the relative entropy $H(\nu|\mu)$ is the analogous of the Guerra-Morato action associated with the free euclidean particle (or field). Furthermore as it is stressed in **[16]** and **[17]** the formula (2.19) is also related to the large deviation theory through Sanov's theorem which yields a very concrete intuition of the experiment. Since the reader may be not familiar with these notions, it seams necessary to recall here briefly and formally the main lines of this stochastic picture of euclidean quantum mechanics. Let $\mathcal{V} : S \to \mathbb{R}$ be a smooth potential which is such that

$$\frac{d\mu_{\mathcal{V}}}{d\mu} := e^{-\int_0^1 \mathcal{V}(W_s) ds}$$

defines a probability equivalent to μ. Jensen's inequality easily implies that the infinimum of

$$\{H(\nu|\mu_{\mathcal{V}})|W_1\nu = \widehat{\nu}\}$$

is attained by the probability ν defined by

$$\frac{d\nu}{d\mu_{\mathcal{V}}} = \frac{1}{E_\mu\left[\frac{d\mu_{\mathcal{V}}}{d\mu}\middle|\sigma(W_1)\right]}\frac{d\widehat{\nu}}{d\widehat{\mu}} \circ W_1$$

Let

$$H = -\frac{\Delta}{2} + \mathcal{V}$$

Then at least formally in the case $\dim S < \infty$, by noting λ the Lebesgue measure on S, it is straightforward to check that we have

$$\frac{dW_t\nu}{d\lambda} = \theta(t,x)\theta^\star(t,x) \tag{2.21}$$

where θ solves

$$\partial_t\theta = H\theta$$

and where $\theta^\star$ is the fundamental solution of the time reversed equation

$$-\partial_t\theta^\star = H\theta^\star$$

By subsituting formally $t \to it$ (this procedure is usually called the rotation of Wick) we would have density $\rho(t,x)$ with the shape

$$\rho(t,x) = \Psi(t,x)\Psi^\star(t,x)$$

where Ψ (resp. $\Psi^\star$) solves the Schrödinger equation (its conjugate). For that reason ν is said to model time imaginary quantum mechanics. By considering the associated spacetime metric, it is also called euclidean quantum mechanics. Note that within this framework, ν is the law of a solution to

$$dX_t = dW_t - \dot{v}_t \circ X dt; X_0 = 0$$

where v is the Girsanov drift associated with ν. The possibility to deal with euclidean quantum mechanics through stochastic mechanics in such a way was first showed in **[61]**. For that reason it seams relevant to call the Girsanov shift $V := I_W + v$ associated with the above measure ν, the Schrödinger shifts associated with $\widehat{\nu}$ under the potential $\mathcal{V}$. Although our results and methods may extend to the case of potentials, we prefered to focus on the free case (i.e. $\mathcal{V} = 0$), and to treat both the finite and the infinite dimensional cases. Henceforth we allow S to be of infinite dimensions, unless otherwise stated. In the next section we will also see that these processes are also involved in information theory. Hence these processes are involved in several fields in which it would be relevant to prove the pathwise uniqueness. For instance in the point of view of stochastic mechanics, the pathwise uniqueness for (2.20) means that the stochastic description of free euclidean quantum mechanics fits with the classical picture of determinism. This is the main motivation of this section in which we give some (very large) sufficient conditions for the Girsanov shift associated with such probabilities to be invertible. This motivates the following definition :

Definition IV.2. *Let $\widehat{\nu}$ be a probability such that $\widehat{\nu} << \widehat{\mu}$. Further note ν the probability on W which is defined by*

$$\frac{d\nu}{d\mu} = \frac{d\widehat{\nu}}{d\widehat{\mu}} \circ W_1$$

and note v the Girsanov drift associated with ν. We say that ν is the (optimal) measure associated with $\widehat{\nu}$ (on the space on the path) and that $V := I_W + v$ (resp. v) is the (free) Schrödinger shift (resp. drift) associated with the final marginal $\widehat{\nu}$.

We will show below that when $\dim S < \infty$ (resp. $\dim S = \infty$) free Schrödinger shifts with a final marginal of finite entropy (resp. with a bounded density with respect to the Wiener measure on S) are always invertible. In particular we don't assume any regularity conditions on the density.

Theorem IV.2. *Let $S = \mathbb{R}^d$ and let $\widehat{\nu}$ be a probability equivalent to the Gaussian measure $\widehat{\mu}$ on $\mathbb{R}^d$ with finite entropy ($H(\widehat{\nu}|\widehat{\mu}) < \infty$). Then the free Schrödinger shift with final marginal $\widehat{\nu}$ is invertible.*

Proof: Let ν be the optimal measure associated with $\widehat{\nu}$ by Definition IV.2 and let v be the Girsanov drift associated with ν so that $V := I_W + v$ is the Schrödinger shift associated with $\widehat{\nu}$. It is well known (see[**16**]) and straightforward to see that the Itô formula yields the following expression for the Schrödinger drift v. For each $t \in [0,1)$

$$\dot{v}_t(\omega) := b(t, W_t) \tag{2.22}$$

where $b : (t,x) \in [0,1) \times \mathbb{R}^d \to -D \ln P_{1-t} \frac{d\widehat{\nu}}{d\widehat{\mu}}(x)$ and where $(t,x) \in [0,1) \times \mathbb{R}^d \to P_{1-t} \frac{d\widehat{\nu}}{d\widehat{\mu}}(x)$ is the heat kernel defined by

$$P_{1-t} \frac{d\widehat{\nu}}{d\widehat{\mu}}(x) = \int_W d\mu(\omega) \frac{d\widehat{\nu}}{d\widehat{\mu}}(W_{1-t} + x)$$

We will show that the finite entropy condition implies that v is locally bounded (see Definition IV.1) so that the shift V will be invertible. First note that for each integer $i \in [1,d]$ $(\dot{v}_s^i, s \in [0,1))$ is a $(\mathcal{F}_t)_{t \in [0,1)}$ martingale on $(W, \mathcal{F}, \nu)$. Indeed for each $t < 1$ and $s \leq t$, let $\theta_s \in C_b(W)$ be $\mathcal{F}_s$ measurable. We then have

$$\begin{aligned}
E_\nu\left[-\dot{v}_t^i \theta_s\right] &= E_\nu\left[D^i \ln P_{1-t} \frac{d\widehat{\nu}}{d\widehat{\mu}}(W_t)\theta_s\right] \\
&= E_\mu\left[\frac{d\nu}{d\mu} D^i \ln P_{1-t} \frac{d\widehat{\nu}}{d\widehat{\mu}}(W_t)\theta_s\right] \\
&= E_\mu\left[E_\mu\left[\frac{d\nu}{d\mu}\middle|\mathcal{F}_t\right] D^i \ln P_{1-t} \frac{d\widehat{\nu}}{d\widehat{\mu}}(W_t)\theta_s\right] \\
&= E_\mu\left[P_{1-t} \frac{d\widehat{\nu}}{d\widehat{\mu}}(W_t) D^i \ln P_{1-t} \frac{d\widehat{\nu}}{d\widehat{\mu}}(W_t)\theta_s\right] \\
&= E_\mu\left[D^i P_{1-t} \frac{d\widehat{\nu}}{d\widehat{\mu}}(W_t)\theta_s\right] \\
&= E_\mu\left[P_{t-s} D^i P_{1-s} \frac{d\widehat{\nu}}{d\widehat{\mu}}(W_s)\theta_s\right] \\
&= E_\mu\left[D^i P_{1-s} \frac{d\widehat{\nu}}{d\widehat{\mu}}(W_s)\theta_s\right] \\
&= E_\nu\left[D^i \ln P_{1-s} \frac{d\widehat{\nu}}{d\widehat{\mu}}(W_s)\theta_s\right] \\
&= E_\nu\left[-\dot{v}_s^i \theta_s\right]
\end{aligned}$$

Hence $t \in [0,1) \to E_\nu\left[(\dot{v}_t^i)^2\right]$ is increasing. Together with the finite energy condition $v \in L_a^2(\nu, H)$ it yields $\dot{v}_t^i \in L^2(\nu)$ for each $t \in [0,1)$. Therefore by Doob's inequality we get that for each $t \in [0,1)$

$$\sup_{s \in [0,t]} |\dot{v}_s|^2_{\mathbb{R}^d} \in L^1(\nu) \tag{2.23}$$

Since $\nu \sim \mu$ equation (2.23) implies that the hypothesis of Theorem IV.1 are satisfied with $\tau_n := 1 - 1/n$. Therefore V is invertible □

In the case where $\dim S = \infty$ it is harder to get a clean expression of the Schrödinger shift since we need an Itô formula. To avoid the use of such a formula we give an elementary proof in Proposition IV.2 which is in the spirit of the proof of the Clark-Ocone formula.

Proposition IV.2. *Let $\widehat{\nu}$ be a probability such that $\widehat{\nu} \sim \widehat{\mu}$ and $\frac{d\widehat{\nu}}{d\widehat{\mu}} \in L^2(\widehat{\mu})$. We further note v the free Schrödinger drift associated with $\widehat{\nu}$ (see Definition IV.2). We then have for each $t \in [0,1)$*

$$\dot{v}_t = b(t, W_t) \tag{2.24}$$

$d\mu \times dt$ a.s. where

$$b(t,x) := -D \ln Q_{1-t} \frac{d\widehat{\nu}}{d\widehat{\mu}}(x)$$

and where Q_t is the heat kernel on S i.e. $Q_t \frac{d\widehat{\nu}}{d\widehat{\mu}}(x) = E_\mu\left[\frac{d\widehat{\nu}}{d\widehat{\mu}}(x+W_t)\right]$ for each $x \in S$.

Proof: Let ν be the measure associated with $\widehat{\nu}$ by Definition IV.2 and let v be the Girsanov drift associated with ν which is also the Schrödinger drift of $\widehat{\nu}$. When $\frac{d\widehat{\nu}}{d\widehat{\mu}} \in L^2(\widehat{\mu})$ the same proof as Lemma 3.3.2 of **[50]** applies and we know that for each $t < 1$ there is a modification of $Q_{1-t}\frac{d\widehat{\nu}}{d\widehat{\mu}}$ such that for each $\omega \in S$ the map $h \in H_S \to Q_{1-t}\frac{d\widehat{\nu}}{d\widehat{\mu}}(\omega + h)$ is real analytic on H_S. We then chose this modification. Hence, it is straightforward to check that for any $k \in H_S$ $Q_t D_k Q_s = D_k Q_t Q_s = D_k Q_{t+s}$ and that $E_\mu\left[\frac{d\widehat{\nu}}{d\widehat{\mu}} \circ W_1 \Big| \mathcal{F}_t\right] = Q_{1-t}\frac{d\widehat{\nu}}{d\widehat{\mu}}(W_t)$. In order to avoid the use of the theory of Watanabe distributions, or of an Itô formula on abstract Wiener space, we give an elementary proof. For any $t < 1$ let $L^2_a(\mu, H, t)$ be the subset of the $u \in L^2_a(\mu, H)$ for which the support of $s \in [0,1] \to \dot{u}_s \in H_S$ is in $[0,t] \subset [0,1)$. We recall that from the martingale representation theorem which holds on W (see **[50]** chap.2), $\left\{\delta^W \alpha, \alpha \in L^2_a(\mu, H, t)\right\}$ is dense in $\left\{X - E_\mu[X] : X \in L^2(\mu)\ ,\ X \textit{ is } \mathcal{F}_t \textit{ measurable}\right\}$. Then we have for each $h \in L^2_a(\mu, H, t)$

$$\begin{aligned}
E_\mu\left[\frac{d\nu}{d\mu} \circ \tau_h \rho(-\delta h)\right] &= E_\mu\left[E_\mu\left[\frac{d\nu}{d\mu} \circ \tau_h \Big| \mathcal{F}_t\right] \rho(-\delta h)\right] \\
&= E_\mu\left[E_\mu\left[\frac{d\widehat{\nu}}{d\widehat{\mu}} \circ (W_1 + h_t) \Big| \mathcal{F}_t\right] \rho(-\delta h)\right] \\
&= E_\mu\left[(Q_{1-t}\frac{d\widehat{\nu}}{d\widehat{\mu}})(W_t + h_t)\rho(-\delta h)\right]
\end{aligned}$$

Hence

$$\begin{aligned}
\frac{d}{d\lambda} E_\mu\left[\left[\frac{d\nu}{d\mu} \circ \tau_{\lambda h} - 1\right] \rho(-\delta \lambda h)\right] |_{\lambda=0} &= \frac{d}{d\lambda} E_\mu\left[\left[(Q_{1-t}\frac{d\widehat{\nu}}{d\widehat{\mu}})(W_t + \lambda h_t) - 1\right] \rho(-\delta\lambda h)\right] |_{\lambda=0} \\
&= E_\mu\left[D_{h_t} Q_{1-t}\frac{d\widehat{\nu}}{d\widehat{\mu}} \circ (W_t)\right] - E_\mu\left[\left[Q_{1-t}\frac{d\widehat{\nu}}{d\widehat{\mu}}(W_t) - 1\right] \delta h\right]
\end{aligned}$$

On the other hand the Cameron-Martin theorem yields

$$E_\mu\left[\left[\frac{d\nu}{d\mu} \circ \tau_{\lambda h} - 1\right] \rho(-\delta\lambda h)\right] = 0$$

By setting

$$\alpha^t_\cdot = \int_0^\cdot 1_{s \le t} D Q_{1-s}\left[\frac{d\widehat{\nu}}{d\widehat{\mu}}\right](W_s) ds$$

we obtain

$$\begin{aligned}
E_\mu\left[(E_\mu\left[\frac{d\nu}{d\mu}\Big|\mathcal{F}_t\right]-1)\delta h\right] &= E_\mu\left[\left[Q_{1-t}\frac{d\widehat{\nu}}{d\widehat{\mu}}(W_t)-1\right]\delta h\right]\\
&= E_\mu\left[< DQ_{1-t}\frac{d\widehat{\nu}}{d\widehat{\mu}}(W_t), h_t >_{H_S}\right]\\
&= \int_0^t E_\mu\left[< DQ_{1-t}\frac{d\widehat{\nu}}{d\widehat{\mu}}(W_t), \dot{h}_s >_{H_S}\right]ds\\
&= \int_0^t E_\mu\left[D_{\dot{h}_s}Q_{1-t}\frac{d\widehat{\nu}}{d\widehat{\mu}}(W_t)\right]ds\\
&= \int_0^t E_\mu\left[E_\mu\left[D_{\dot{h}_s}Q_{1-t}\frac{d\widehat{\nu}}{d\widehat{\mu}}(W_t)\Big|\mathcal{F}_s\right]\right]ds\\
&= \int_0^t E_\mu\left[Q_{t-s}D_{\dot{h}_s}Q_{1-t}\frac{d\widehat{\nu}}{d\widehat{\mu}}(W_s)\right]ds\\
&= \int_0^t E_\mu\left[D_{\dot{h}_s}Q_{1-s}\frac{d\widehat{\nu}}{d\widehat{\mu}}(W_s)\right]ds\\
&= \int_0^t E_\mu\left[< DQ_{1-s}\frac{d\widehat{\nu}}{d\widehat{\mu}}\circ(W_s), \dot{h}_s >_{H_S}\right]ds\\
&= E_\mu\left[< \alpha^t, h >_H\right]\\
&= E_\mu\left[\delta^W\alpha^t\delta^W h\right]
\end{aligned}$$

which means that $E_\mu\left[\frac{d\nu}{d\mu}\Big|\mathcal{F}_t\right]-1=\delta^W\alpha^t$ i.e.

$$E_\mu\left[\frac{d\nu}{d\mu}\Big|\mathcal{F}_t\right]=1+\int_0^t DQ_{1-s}\frac{d\widehat{\nu}}{d\widehat{\mu}}\circ(W_s)dW_s$$

by construction of v (see Section 2) it yields (2.24) □

Lemma IV.1. *Let $L\in L^\infty(\widehat{\mu})$, $\epsilon\in(0,1)$ and $T\in[0,1)$. Let $b:[0,1]\times S\to H_S$ be the mapping defined by*

$$b(t,x):=-D\ln P_{1-t}L(x)1_{t<T}1_{P_{1-t}L(x)>\epsilon}$$

and let v be defined by

$$v:=\int_0^\cdot b(s,W_s)ds$$

Then $V:=I_W+v$ is invertible.

Proof: For any $h\in H_S$, $x\in S$ and $t<1$ we have (see **[25]**)

$$\begin{aligned}
|D_h\ln Q_{1-t}L(x)| &= \left|\frac{1}{\sqrt{1-t}}\frac{\int_S(< h,y >)L(x+\sqrt{1-t}y)\widehat{\mu}(dy)}{\int_S L(x+\sqrt{1-t}y)\widehat{\mu}(dy)}\right|\\
&\leq \frac{|h|_H K}{\sqrt{1-t}P_{1-t}L(x)}
\end{aligned}$$

where K is the essential supremum of L. Therefore $|D\ln Q_{1-t}L(x)|_H\leq\frac{K}{\sqrt{1-t}P_{1-t}L(x)}$ and

$$|b(t,x)|=|D\ln P_{1-t}L(x)1_{t<T}1_{P_{1-t}L(x)>\epsilon}|\leq\frac{K}{\epsilon\sqrt{1-T}}$$

In particular $v\in L^0_a(\mu,H)$ and together with Novikov's criterion it yields $E_\mu\left[\rho\left(-\delta^W v\right)\right]=1$. Moreover if we assume that V is right invertible with an inverse $U:=I_W+u$, the boundedness of b will imply $E_\mu\left[\rho(-\delta^W u)\right]=1$ so that the invertibility will automatically follow from Proposition II.1. Hence to get the invertibility of V, it suffices to prove that V is right invertible. Specifically, since b is bounded it suffices to prove that

$$dX_t = dB_t - \dot{v}_t \circ X dt \tag{2.25}$$

has a unique strong solution. Since S is a Polish space the Yamada-Watanabe criterion (see **[22]**) also applies for (2.25). Therefore to prove the existence of a strong solution it suffices to check that a weak solution of equation (2.25) exists, and that the pathwise uniqueness holds. Since b is bounded, we already have the existence of a weak solution and the uniqueness in law for (2.25) by transformation of the drift (see**[22]**). Hence we only have to prove the pathwise uniqueness. Before we do this, we have make some preparations and to prove that $\mu - a.s.$ for any $t < 1$ and $(h, k) \in \mathcal{H}_{t,\omega} \times \mathcal{H}_{t,\omega}$ we have

$$|D \ln P_{1-t} L(W_t + h) - D \ln P_{1-t} L(W_t + k)|_{H_S} \leq C|h - k|_{H_S} \tag{2.26}$$

for a $C > 0$ which depends on T. For any $(t, \omega) \in [0, 1) \times W$ we set

$$\mathcal{H}_{t,\omega} := \{h \in H_S : P_{1-t} L(W_t + h) > \epsilon\}$$

Since $x \to P_{1-t}L(x)$ is H_S-continuous (end even $H_S - C^\infty$ see **[25]**) $\mathcal{H}_{t,\omega}$ is an open set. On the other hand we have for any $h, k \in H_S$, and $x \in S$

$$D^2_{h,k} \ln Q_{1-t} L(x) = \frac{D_h D_k Q_{1-t} L}{Q_{1-t} L} - \frac{(D_h Q_{1-t} L)(D_k Q_{1-t} L)}{(Q_{1-t} L)^2}$$

and

$$\begin{aligned} \left|D^2_{h,k} Q_{1-t} L(x)\right| &= \frac{1}{1-t} \left| \int (< h, y >< k, y > - < h, k >_{H_S}) L(x + \sqrt{1-t} y) \widehat{\mu}(dy) \right| \\ &\leq \frac{K}{1-t} |h|_{H_S} |k|_{H_S} \end{aligned}$$

where the last inequality follows from

$$\int (< h, y >< k, y > - < h, k >_{H_S})^2 \widehat{\mu}(dy) = |h|^2_{H_S} |k|^2_{H_S}$$

Hence by taking $C := \frac{K}{(1-T)\epsilon} + (\frac{K}{\sqrt{1-T}\epsilon})^2$, we get (2.26). We can now begin the proof of pathwise uniqueness. Let U and $\widetilde{U}$ be two solutions with initial distribution α, which are defined on a same space $(\Omega, \mathcal{G}, \mathbb{P})$, with a same filtration $(\mathcal{G}_t)$ defined on it, and with a same $(\mathcal{G}_t)-$Brownian motion(B_t). We define two $(\mathcal{G}_t)-$optional times :

$$\tau := \inf (\{t : |u_t - \widetilde{u}_t|_{H_S} > 0\}) \wedge 1 \tag{2.27}$$

and

$$\sigma := \inf \left(\left\{ t : (P_{1-t} L \circ U_t - \epsilon)(P_{1-t} L \circ \widetilde{U}_t - \epsilon) < 0 \right\} \right) \wedge 1 \tag{2.28}$$

Let $\widetilde{\Omega}$ be the set of the $\omega \in \Omega$ for which both $t \to U_t$ and $t \to \widetilde{U}_t$ are continuous. By construction it is such that $\mathbb{P}(\widetilde{\Omega}) = 1$. Moreover for any $\omega \in \widetilde{\Omega}$ the continuity implies that we have $\sigma(\omega) \leq \tau(\omega)$ if and only if $\tau(\omega) = 1$. Therefore, to prove pathwise uniqueness it suffices prove that $\mu - a.s.$ $\sup_{t \in [0, \sigma(\omega)]} |u_t - \widetilde{u}_t|_{H_S} = 0$. For any $t \in [0, \sigma(\omega)]$ the definition of σ implies :

$$\begin{aligned} |u_t - \widetilde{u}_t|_{H_S} &\leq \int_0^t |b(s, U_s) - b(s, \widetilde{U}_s)| ds \\ &\leq A + B \end{aligned}$$

where

$$A := \int_0^t |b(s, U_s) - b(s, \widetilde{U}_s)| 1_{\mathcal{H}_{s,\omega} \times \mathcal{H}_{s,\omega}}(u_s, \widetilde{u}_s)$$

and

$$B := \int_0^t |b(s, U_s) - b(s, \widetilde{U}_s)| 1_{(\mathcal{H}_{s,\omega})^c \times (\mathcal{H}_{s,\omega})^c}(u_s, \widetilde{u}_s)$$

and where $(\mathcal{H}_{s,\omega})^c$ is the complement of $\mathcal{H}_{s,\omega}$ in H_S. From the definitions of $\mathcal{H}_{t,\omega}$ and of $b(t,x)$ $\mu - a.s.$ $B = 0$. On the other hand (2.26) implies

$$\begin{aligned} A &\leq C \int_0^t |u_s - \widetilde{u}_s|_{H_S} 1_{\mathcal{H}_{s,\omega} \times \mathcal{H}_{s,\omega}}(u_s, \widetilde{u}_s) \\ &\leq C \int_0^t |u_s - \widetilde{u}_s|_{H_S} \end{aligned}$$

Hence, up to a a negligible set, Gronwall's lemma implies that $u_t(\omega) = \widetilde{u}_t(\omega)$ for every $t < \sigma(\omega)$. This proves that $\mu - a.s.$ $\sigma \leq \tau$, from which we know that $\mu - a.s.$ $\tau = 1$. □

Theorem IV.3. *Let $\widetilde{\nu}$ be a probability equivalent to $\widehat{\mu}$ such that $\frac{d\widehat{\nu}}{d\widehat{\mu}} \in L^\infty(\mu)$. Then the free Schrödinger shift with final marginal $\widehat{\nu}$ is invertible.*

Proof: Let v be the Girsanov drift associated with the probability $\nu := \frac{d\widehat{\nu}}{d\widehat{\mu}} \circ W_1$. From Definition IV.2 $V := I_W + v$ is the free Schrödinger shift associated with $\widehat{\nu}$. For any $(n,T) \in \mathbb{N} \times [0,1)$ we set

$$v^{n,T} := \int_0^{\cdot} b^{n,T}(s, W_s) ds$$

where

$$b^{n,T}(t,x) = -D \ln P_{1-t} \frac{d\widehat{\nu}}{d\widehat{\mu}}(x) 1_{t \leq T} 1_{P_{1-t}\frac{d\widehat{\nu}}{d\widehat{\mu}}(x) > \frac{1}{n}}$$

and

$$v^T := -\int_0^{\cdot} D \ln P_{1-t} \frac{d\widehat{\nu}}{d\widehat{\mu}}(x) 1_{t \leq T} dt$$

By Lemma IV.1 all the $V^{n,T} := I_W + v^{n,T}$ are invertible. Let $\tau_n := \inf(\{t : P_{1-t}L(W_t) < \frac{1}{n}\} \wedge 1$. Since $\widehat{\nu} \sim \widehat{\mu}$ we have $\mu - a.s.$ $(\tau_n) \uparrow 1$. On the other hand from the definitions $\pi_{\tau_n} v^T = \pi_{\tau_n} v^{n,T}$. From Corollary III.1 it implies that V^T is invertible for any $T < 1$. Taking $\tau_n := 1 - \frac{1}{n}$ yields the local invertibility of V which is therefore invertible (Theorem III.2). □

3. Invertibility of the free Schrödinger shifts in the perspective of optimal transport

In this section we will handle the following sets

Definition IV.3. *Let $\widehat{\nu}$ be a probability equivalent to $\widehat{\mu}$ then we set*

$$\mathcal{R}(\mu, \widehat{\nu}) := \{u \in L^0(\mu, H) : U_1\mu = \widehat{\nu} \text{ where } U := I_W + u\}$$

and

$$\mathcal{R}_a(\mu, \widehat{\nu}) := L^0_a(\mu, H) \cap \mathcal{R}(\mu, \widehat{\nu})$$

In Section 2 we have given mild sufficient conditions for the invertibility of free Schrödinger shifts. It also involves a generalization for the representation formula of **[26]** which is given in Proposition IV.3. We recall that in **[26]** the Proposition IV.3 is proved on the classical Wiener space under the condition that $\frac{d\widehat{\nu}}{d\widehat{\mu}} = e^f$ where f is C^2 with all its derivatives bounded. Moreover this latter does not relate the equality case to invertibility but rather to right invertibility.

Proposition IV.3. *Let $\widehat{\nu}$ be a probability equivalent to $\widehat{\mu}$ with finite entropy i.e. $H(\widehat{\nu}|\widehat{\mu}) < \infty$. If $\dim S = \infty$ further assume that $\frac{d\widehat{\nu}}{d\widehat{\mu}} \in L^\infty(\widehat{\mu})$. Then we have*

$$2H(\widehat{\nu}|\widehat{\mu}) = \min\left(\left\{E_\mu\left[|u|_H^2\right] : u \in \mathcal{R}_a(\mu, \widehat{\nu})\right\}\right) \tag{3.29}$$

Moreover the infimum is attained by a $U := I_W + u$ which is the inverse of the Schrödinger shift (see Definition IV.2) associated with $\widehat{\nu}$.

Proof: Let $\widehat{\nu}$ be a probability equivalent to $\widehat{\mu}$. From equation (2.19) for any $u \in \mathcal{R}(\mu, \widehat{\nu})$ we have

$$H(\widehat{\nu}|\widehat{\mu}) \leq H(U\mu|\mu) \tag{3.30}$$

Then Theorem V.2 implies that

$$2H(U\mu|\mu) \leq E_\mu\left[|u|_H^2\right] \tag{3.31}$$

Together with (3.30) it yields the inequality in (3.29). Moreover from formula (2.19) the equality in (3.30) is attained by the optimal measure associated with ν. By applying Theorem IV.2 (or Theorem IV.3 if $\dim S = \infty$) we get the existence of a u which attains the equality in (3.31). Hence the inverse of the free Schrödinger shift associated with $\widehat{\nu}$ attains the optimum in (3.29). □

Naturally related questions are whether a similar representation formula holds for the Wasserstein distance $\sqrt{d(\widehat{\nu}, \widehat{\mu})}$ and whether the infimum is also attained by an invertible shift T with a given marginal $\widehat{\nu}$. The answer is given by Proposition IV.4. Also note that the law of the non$-$adapted shift T which appears in the statement of Proposition IV.4 is the analogous in the non-adapted case to the optimal law $\nu := \frac{d\widehat{\nu}}{d\widehat{\mu}} \circ W_1$.

Proposition IV.4. *Let $\widehat{\nu}$ be a probability equivalent to $\widehat{\mu}$ of finite Wasserstein distance (see Definition I.1) with respect to $\widehat{\mu}$. Then $d(\widehat{\nu}, \widehat{\mu})$ is given by*

$$d(\widehat{\nu}, \widehat{\mu}) = \inf\left(\left\{E_\mu\left[|u|_H^2\right] : u \in \mathcal{R}(\mu, \widehat{\nu})\right\}\right) \tag{3.32}$$

and the infimum is attained by an invertible shift T such that

$$W_1 \circ T = T^S \circ W_1 \tag{3.33}$$

Moreover, let $\widetilde{S}$ be the inverse of T i.e. $\mu - a.s.$ $T \circ S = I_W$ and $S \circ T = I_W$. We also have

$$W_1 \circ \widetilde{S} = \widetilde{S}^S \circ W_1 \tag{3.34}$$

where T^S and $\widetilde{S}^S$ are the solutions of the Monge problem (resp. its inverse) on S defined in Theorem VI.2

Proof: We recall that I_S denotes the identity map on S. We note $\mathcal{R}(\widehat{\mu}, \widehat{\nu})$ the set of the mappings $u^S \in L^2(\widehat{\mu}, H_S)$ such that $U^S\widehat{\mu} = \widehat{\nu}$ where $U^S = I_S + u^S$. It may be seen as a subset of $\mathcal{R}(\mu, \widehat{\nu})$. Indeed for any such u^S we can set $i(u^S)_. = \int_0^. u^S \circ W_1 ds$. Obviously $i(u^s) \in \mathcal{R}(\mu, \widehat{\nu})$ so that $i(\mathcal{R}(\widehat{\mu}, \widehat{\nu})) \subset \mathcal{R}(\mu, \widehat{\nu})$ and

$$\begin{aligned}\inf\left(\left\{E_\mu\left[|i(u^S)|_H^2\right] : u^S \in \mathcal{R}(\widehat{\mu}, \widehat{\nu})\right\}\right) &= \inf\left(\left\{E_\mu\left[|u|_H^2\right] : u \in i(\mathcal{R}(\widehat{\mu}, \widehat{\nu}))\right\}\right) \\ &\geq \inf\left(\left\{E_\mu\left[|u|_H^2\right] : u \in \mathcal{R}(\mu, \widehat{\nu})\right\}\right)\end{aligned}$$

Of course we have

$$E_\mu\left[\left|i(u^S)\right|_H^2\right] = E_{\widehat{\mu}}\left[\left|u^S\right|_{H_S}^2\right] \tag{3.35}$$

and

$$\begin{aligned}d(\widehat{\nu}, \widehat{\mu}) &= \inf\left(\left\{E_{\widehat{\mu}}\left[|u^S|_{H_S}^2\right] : u^S \in \mathcal{R}(\widehat{\mu}, \widehat{\nu})\right\}\right) \\ &= \inf\left(\left\{E_\mu\left[|i(u^S)|_H^2\right] : u^S \in \mathcal{R}(\widehat{\mu}, \widehat{\nu})\right\}\right)\end{aligned}$$

Hence we get

$$d(\widehat{\nu},\widehat{\mu}) \geq \inf\left(\left\{E_\mu\left[|u|_H^2\right] : u \in \mathcal{R}(\mu,\widehat{\nu})\right\}\right) \tag{3.36}$$

On the other hand for any $u \in \mathcal{R}(\mu,\widehat{\nu})$ Jensen's inequality yields

$$E_\mu\left[|u_1|_{H_S}^2\right] \leq E_\mu\left[|u|_H^2\right] \tag{3.37}$$

and since $(U_1 \times W_1)\mu \in \Sigma(\widehat{\nu},\widehat{\mu})$ we have

$$d(\widehat{\nu},\widehat{\mu}) \leq E_\mu\left[|u_1|_{H_S}^2\right] \tag{3.38}$$

The inequalities (3.37) and (3.38) clearly yield

$$d(\widehat{\nu},\widehat{\mu}) \leq \inf\left(\left\{E_\mu\left[|u|_H^2\right] : u \in \mathcal{R}(\mu,\widehat{\nu})\right\}\right) \tag{3.39}$$

Together with (3.36) the inequality (3.39) yields (3.32). Now we set $T := I_W + i(t^S)$ where $t^S := T^S - I_S$ i.e.

$$T : (\sigma,\omega) \in [0,1] \times W \to T_\sigma := W_\sigma + \sigma t^S \circ W_1$$

and we want to show that it attains the infimum of (3.32). Indeed by hypothesis we have (3.33) so that $T_1\mu = T^S\widehat{\mu} = \widehat{\nu}$. Hence

$$t := T - I_W \in \mathcal{R}(\mu,\widehat{\nu})$$

On the other hand (3.35) yields

$$\begin{aligned} E_\mu\left[|T - I_W|_H^2\right] &= E_{\widehat{\mu}}\left[|T^S - I_S|_{H_S}^2\right] \\ &= d(\widehat{\nu},\widehat{\mu}) \end{aligned}$$

Hence t attains the infimum of (3.32). We now prove the last part of the claim. It is easy to see that if we set $\widetilde{S}_\sigma := W_\sigma + \sigma\widetilde{s}^S \circ W_1$ where $\widetilde{s}^S := \widetilde{S}^S - I_S$ we then have

$$\begin{aligned} (T \circ \widetilde{S})_\sigma &= W_\sigma + \sigma\widetilde{s}^S \circ W_1 + \sigma T^S \circ (W_1 + \widetilde{s}^S \circ W_1) \\ &= W_\sigma + \sigma(T^S \circ \widetilde{S}^S - I_S) \circ W_1 \\ &= W_\sigma \end{aligned}$$

Which shows that $T := I_W + t$ is invertible with inverse $\widetilde{S}$. □

Under the hypothesis of Proposition IV.3, we then have

$$d(\widehat{\mu},\widehat{\nu}) = \inf\left(\left\{E_\mu\left[|u|_H^2\right] : u \in \mathcal{R}(\mu,\widehat{\nu})\right\}\right)$$

and

$$2H(\widehat{\nu},\widehat{\mu}) = \inf\left(\left\{E_\mu\left[|u|_H^2\right] : u \in \mathcal{R}_a(\mu,\widehat{\nu})\right\}\right)$$

Since

$$\mathcal{R}_a(\mu,\widehat{\nu}) \subset \mathcal{R}(\mu,\widehat{\nu})$$

we get again the Talagrand inequality. Moreover in the proof of Proposition IV.4 the existence of T clearly follows from the existence of an invertible shift on S which solves the Monge problem. This suggests to investigate the connection between the invertibility of the Schrödinger shifts and the problem of invertibility on S. For that reason henceforth and until the end of this section we assume that $S = C_0([0,1],\mathbb{R}^d)$ (however our results extend to the case where S is an abstract Wiener space with a time structure as in [**49**]). Let $\widehat{\nu}$ be a probability equivalent to $\widehat{\mu}$, we then have the existence of a Girsanov shift $V^S := I_S + v^S$ such that $\widehat{\mu} - a.s.$

$$\frac{d\widehat{\nu}}{d\widehat{\mu}} = \exp\left(-\delta^{W^S} v^S - \frac{|v^S|_{H_S}}{2}\right)$$

and such that $t \to V_t^S$ is a Wiener process under $\widehat{\nu}$ on S. We recall that in that case $\delta^{W^S} v^S$ denotes the stochastic integral of v^S with respect to the coordinate process $t \to W_t^S$ on S. $L_a^0(\widehat{\mu}, H_S)$ is the subset of the elements of $L^0(\widehat{\mu}, H_S)$ which are adapted to the filtration generated by the coordinate process on S . We call v^S (resp. $V^S := I_S + v^S$ where I_S is the identity map on S) the $S-$Girsanov drift (resp. shift) associated with $\widehat{\nu}$. In this case we also define

$$\mathcal{R}(\widehat{\mu}, \widehat{\nu}) := \left\{u \in L^0(\mu, H_S) : U\widehat{\mu} = \widehat{\nu} \text{ where } U := I_S + u\right\}$$

and $\mathcal{R}_a(\widehat{\mu}, \widehat{\nu}) = \mathcal{R}(\widehat{\mu}, \widehat{\nu}) \cap L_a^0(\widehat{\mu}, H)$. Proposition IV.5 and Proposition IV.6 complete the analogy between (3.29) and (3.32) : in particular (3.43) has to be compared with (3.33).

Proposition IV.5. *Let $\widehat{\nu}$ be a probability equivalent to $\widehat{\mu}$ on $S = C_0([0,1], \mathbb{R}^d)$ which is such that $\frac{d\widehat{\nu}}{d\widehat{\mu}} \in L^\infty(\widehat{\mu})$, and let $V := I_W + v$ be the Schrödinger shift associated with $\widehat{\nu}$ (see Definition IV.2). Then we have*

$$W_1 \circ V = V^S \circ W_1 \tag{3.40}$$

where $V^S := I_S + v^S$ is the $S-$Girsanov shift associated with $\widehat{\nu}$. If we denote by $U := I_W + u$ the inverse of the Schrödinger shift we also have equivalently

$$V^S \circ U_1 = W_1 \tag{3.41}$$

and

$$E_\mu\left[|u_1|_{H_S}^2\right] = E_\mu\left[|u|_H^2\right]$$

In particular the following variational formula holds

$$\begin{aligned} H(\widehat{\nu}|\widehat{\mu}) &= E_\nu\left[\frac{|v_1|_{H_S}^2}{2}\right] \\ &= \inf\left(\left\{E_\mu\left[\frac{|a_1|_{H_S}^2}{2}\right] : a \in \mathcal{R}_a(\mu, \widehat{\nu})\right\}\right) \end{aligned}$$

Proof: Let $a \in \mathcal{R}_a(\mu, \widehat{\nu})$ and let $\widehat{a_1}^{A_1}$ be the projection of a_1 on the closed subspace

$$\left\{\theta \circ A_1 : \theta \in L_a^2(\widehat{\nu}, H_S)\right\}$$

Since

$$(\delta^{W^S}\theta) \circ A_1 - \delta^{W_1}(\theta \circ A_1) =< \theta \circ A_1, a_1 >_{H_S}$$

we obtain

$$\begin{aligned} E_\mu\left[< a_1, \theta \circ A_1 >_{H_S}\right] &= E_\mu\left[(\delta^{W^S}\theta) \circ A_1\right] - E_\mu\left[\delta^{W_1}(\theta \circ A_1)\right] \\ &= E_{\widehat{\nu}}\left[\delta^{W^S}\theta\right] \\ &= E_{\widehat{\nu}}\left[\delta^{V^S}\theta\right] - E_{\widehat{\nu}}\left[< v^S, \theta >_{H_S}\right] \\ &= -E_\mu\left[< v^S \circ A_1, \theta \circ A_1 >_{H_S}\right] \end{aligned}$$

Hence

$$\widehat{a_1}^{A_1} + v^S \circ A_1 = 0$$

$\widehat{\mu}$ a.s. In particular together with Proposition I.1 it yields

$$\begin{aligned} 2H(\widehat{\nu}|\widehat{\mu}) &= E_{\widehat{\nu}}\left[|v^S|_{H_S}^2\right] \\ &= E_\mu\left[|v^S \circ A_1|_{H_S}^2\right] \\ &= E_\mu\left[|\widehat{a_1}^{A_1}|_{H_S}^2\right] \\ &\leq E_\mu\left[|a_1|_{H_S}^2\right] \end{aligned}$$

Then Jensen's inequality implies

$$2H(\widehat{\nu}|\widehat{\mu}) \leq E_\mu \left[|a|_H^2\right] \tag{3.42}$$

for any $a \in \mathcal{R}_a(\mu, \widehat{\nu})$. By applying this result to the optimal shift $U := I_W + u$ which is the inverse of V and the Theorem V.2 we obtain

$$2H(\widehat{\nu}|\widehat{\mu}) \leq E_\mu \left[|u|_H^2\right] = 2H(\nu|\mu) = 2H(\widehat{\nu}|\widehat{\mu})$$

where the last equality is a consequence of Proposition V.1 and where ν is the optimal probability associated with $\widehat{\nu}$ by Definition IV.2. Hence in that case the inequalities are equalities and $\widehat{u_1}^{U_1} = u_1$ so that $u_1 = -v^S \circ U_1$ and (3.41) is proved. By applying V to both terms of (3.41) we get (3.40). From this the result comes easily. □

Proposition IV.6. *With the same hypothesis and notations as in Proposition IV.5, let $\widehat{\nu}$ be a probability equivalent to $\widehat{\mu}$ such that $\frac{d\widehat{\nu}}{d\widehat{\mu}} \in L^\infty(\widehat{\mu})$. Moreover, let U be the optimal shift given by Proposition IV.3 which is the inverse of the Schrödinger shift associated with $\widehat{\nu}$. Then the following assertions are equivalent*

- *There is a measurable mapping $U^S : S \to S$ such that $\mu - a.s.$*

$$W_1 \circ U = U^S \circ W_1 \tag{3.43}$$

- *There is a $u^S \in \mathcal{R}_a(\widehat{\mu}, \widehat{\nu})$ such that $U^S = I_S + u^S$ is the both sided $\widehat{\mu}$ almost sure inverse of $V^S := I_S + v^S$ where v^S is the $S-$Girsanov drift associated with $\widehat{\nu}$*

Moreover, in that case, both the U^S are the same and

$$2H(\widehat{\nu}|\widehat{\mu}) = E_{\widehat{\mu}} \left[|u^S|_{H_S}^2\right]$$

Proof: Assume that there is a mapping U^S such that (3.43) holds. We have

$$U^S\widehat{\mu} = U^S \circ W_1\mu = U_1\mu = \widehat{\nu}$$

Moreover, by (3.41) of Proposition IV.5 we get

$$\begin{aligned} \widehat{\mu}\left(\left\{\omega \in S : V^S \circ U^S = I_S\right\}\right) &= \mu\left(\left\{\omega \in W : V^S \circ U^S \circ W_1 = W_1\right\}\right) \\ &= \mu\left(\left\{\omega \in W : V^S \circ U_1 = W_1\right\}\right) \\ &= 1 \end{aligned}$$

Hence $V^S \circ U^S = I_S$ $\widehat{\mu}-a.s.$ Since $U^S\widehat{\mu} = \widehat{\nu} \sim \widehat{\mu}$ Proposition II.1 also implies that $\widehat{\mu}-a.s.$ we have $U^S \circ V^S = I_S$. Hence U^S is the both sided inverse of V^S. This implies that $t \to U_t^S$ generates the same filtration as the coordinate process on S. Let $u^S = U^S - I_S$. Since $u^S = -v^S \circ U^S$, $E_{\widehat{\mu}}\left[|u^S|_{H_S}^2\right] = E_\nu\left[|v^S|_H^2\right] = H(\widehat{\nu}|\widehat{\mu})$. Moreover, since $v^S \in L_a^0(\widehat{\mu}, H_S)$ is adapted to the filtration generated by the coordinate process on S, u^S is also adapted to the filtration generated by U^S and hence to the filtration generated by the coordinate process on S, i.e. $u^S \in L_a^0(\widehat{\mu}, H_S)$. Conversely assume that $U^S = I_W + u^S$ where $u^S \in \mathcal{R}_a(\widehat{\mu}, \widehat{\nu})$ is the inverse of V^S. By applying U^S to both terms of formula (3.41) in Proposition IV.5 we finally get (3.43). □

V

Applications to information theory : Information loss on the path space and Shannon's inequality

Abstract. **Talagrand and Sobolev's inequality on abstract Wiener space. Criterion of invertibility expressed with variance. Shannon's inequality on abstract Wiener space. Interpretation of Shannon's inequality as an information loss in Gaussian Channel.**

The idea to use h-path processes in information theory is not new and we found it implicitly in an original but somehow misleading unpublished paper ([**26**]) in the case of the classical Wiener space. We generalized and clarified some of these results. By completing that work we were acquainted that the author of [**26**] also took the same way. Nevertheless our results are still more general and we find it interesting enough to be presented here. Although the essential ideas of the proofs are not new there are several original contributions in this chapter. First we give a version of a Brascamp-Lieb inequality which holds on any abstract Wiener space. Then we give an abstract Wiener space version of Shannon's inequality which holds on any abstract Wiener space. Since the Lebesgue measure is no more well defined when $\dim S = \infty$ we had to write it in terms of Gaussian measure. We succeeded in this task by making a change of measure from which we get a formulation of Shannon's inequality which seems to be new. By making an analogous change of measure on the path space W we show that Üstünel's criterion (the main result of [**47**]) of which we present a slight generalization here (Theorem V.2) may be written in terms of variance, which seems to be new. Within this framework we reduce the proof of [**26**] as a consequence of the additive properties of variance. Moreover our precise results show that Shannon's inequality as well as the Brascamp Lieb inequality may be seen as the result of information loss in a Gaussian channel.

1. Talagrand's inequality and Sobolev's inequality on any abstract Wiener space

The Monge-Kantorovich problem on abstract Wiener space has been investigated for instance in [**12**] or [**13**] (see also [**55**] for a general overview on this topic). The results of this subsection are almost trivial and may be well known. However it seems relevant to give it here for pedagogical reasons. As a matter of fact here are the two simplest cases in which h-path processes can be used to yield inequalities on any abstract Wiener space from inequalities on the path space. We first recall the following result which is a particular case of the Theorem 3.2 of [**13**] in order to achieve the proof of Proposition V.1

Theorem V.1. *Let $\widehat{\nu}$ be a probability such that $\widehat{\nu} << \widehat{\mu}$ (i.e. absolutely continuous). Assume that $d(\widehat{\nu}, \widehat{\mu}) < \infty$. Then there is a measurable mapping $T^S : S \to S$ which is solution to the original Monge problem. Moreover its graph supports the unique solution of the Monge-Kantorovich problem γ i.e.*

$$(I_S \times T^S)\widehat{\mu} = \gamma$$

in particular $T^S\widehat{\mu} = \widehat{\nu}$, $T^S - I_S \in L^2(\widehat{\mu}, H)$, and there is a mapping $\widetilde{S}^S := (T^S)^{-1}$ such that

$$\mu\left(\left\{\omega | \widetilde{S}^S \circ T^S = I_S\right\}\right) = \nu\left(\left\{\omega | T^S \circ \widetilde{S}^S = I_S\right\}\right) = 1$$

The next Proposition sums up basic properties of the optimal measure associated with a marginal.

Proposition V.1. *Let $\widehat{\nu}$ be a probability such that $\widehat{\nu} \sim \widehat{\mu}$ with $H(\widehat{\nu}|\widehat{\mu}) < \infty$. Let ν be the measure associated with $\widehat{\nu}$ by Definition IV.2. We then have*

- $H(\widehat{\nu}|\widehat{\mu}) = H(\nu|\mu)$
- $d(\widehat{\nu}, \widehat{\mu}) \leq d(\nu, \mu)$

where $d(\nu, \mu)$ is given by Definition I.1 If we further assume that $\frac{d\widehat{\nu}}{d\widehat{\mu}} \in \mathbb{D}_{2,1}(\widehat{\mu})$ we have

$$J(\widehat{\nu}, \widehat{\mu}) = J(\nu, \mu)$$

where

$$J(\widehat{\nu}, \widehat{\mu}) := E_{\widehat{\nu}}\left[\left|D \ln \frac{d\widehat{\nu}}{d\widehat{\mu}}\right|^2_{H_S}\right]$$

and where $J(\nu, \mu)$ is defined in Proposition I.3

Proof: Let v be the Girsanov drift associated with ν. The fact that $H(\widehat{\nu}|\widehat{\mu}) = H(\nu|\mu)$ is obvious. Let

$$s \to T_s := W_s + \int_0^s \dot{t}_s ds$$

be the optimum given by Theorem VI.2 which attains $d(\nu, \mu)$. By definition we have

$$d(\nu, \mu) = E_\mu\left[\int_0^1 |\dot{t}_s|^2_{H_S} ds\right]$$

On the other hand

$$\begin{aligned} T_1\mu &= W_1 T\mu \\ &= W_1\nu \\ &= \widehat{\nu} \end{aligned}$$

and $W_1\mu = \widehat{\mu}$ so that :

$$d(\widehat{\nu}, \widehat{\mu}) \leq E_\mu\left[|T_1 - W_1|^2_{H_S}\right]$$

Hence Jensen's inequality implies

$$\begin{aligned} d(\widehat{\nu}, \widehat{\mu}) &\leq E_\mu\left[|T_1 - W_1|^2_{H_S}\right] \\ &\leq E_\mu\left[\left|\int_0^1 \dot{t}_s ds\right|^2_{H_S}\right] \\ &\leq E_\mu\left[\int_0^1 |\dot{t}_s|^2_{H_S} ds\right] \\ &= d(\nu, \mu) \end{aligned}$$

We now state the last part of the claim. From Proposition I.1 we have

$$\begin{aligned} \nabla_h \frac{d\nu}{d\mu} &= \nabla_h \frac{d\widehat{\nu}}{d\widehat{\mu}} \circ W_1 \\ &= < h_1, D\frac{d\widehat{\nu}}{d\widehat{\mu}} \circ W_1 >_{H_S} \\ &= \int_0^1 < \dot{h}_s, D\frac{d\widehat{\nu}}{d\widehat{\mu}} \circ W_1 >_{H_S} ds \end{aligned}$$

so that $D_s \ln \frac{d\nu}{d\mu} = D \ln \frac{d\widehat{\nu}}{d\widehat{\mu}} \circ W_1$. We recalled above that Clark-Ocone formula yields $\dot{v}_1 = D_1 \ln \frac{d\nu}{d\mu}$. However in that case it can also be seen directly with the same proof as Proposition IV.2, but in that case with $t = 1$.

Since $W_1\nu = \widehat{\nu}$ we get

$$\begin{aligned}
J(\nu,\mu) &= E_\nu\left[|\nabla \ln \frac{d\nu}{d\mu}|_H^2\right] \\
&= \int_0^1 E_\nu\left[\left|D_s \ln \frac{d\nu}{d\mu}\right|_{H_S}^2\right] ds \\
&= \int_0^1 E_\nu\left[\left|D \ln \frac{d\widehat{\nu}}{d\widehat{\mu}} \circ W_1\right|_{H_S}^2\right] ds \\
&= E_\nu\left[\left|D \ln \frac{d\widehat{\nu}}{d\widehat{\mu}} \circ W_1\right|_{H_S}^2\right] \\
&= E_{\widehat{\nu}}\left[\left|D \ln \frac{d\widehat{\nu}}{d\widehat{\mu}}\right|_{H_S}^2\right] \\
&= J(\widehat{\nu},\widehat{\mu})
\end{aligned}$$

□

Proposition V.2 shows how we can use h-path processes to get inequalities on S from inequalities on the path space W.

Proposition V.2. *Let $\widehat{\nu}$ be a probability equivalent to $\widehat{\mu}$ with $H(\widehat{\nu}|\widehat{\mu}) < \infty$. Let $d(\widehat{\mu},\widehat{\nu})$ be the Wasserstein distance, then we have*

$$d(\widehat{\mu},\widehat{\nu}) \leq 2H(\widehat{\nu}|\widehat{\mu})$$

Assume henceforth that $\frac{d\widehat{\nu}}{d\widehat{\mu}} \in \mathbb{D}_{2,1}(\widehat{\mu})$ and let $J(\widehat{\nu}|\widehat{\mu})$ be the Fisher information, which is defined by

$$J(\widehat{\nu}|\widehat{\mu}) = E_{\widehat{\nu}}\left[\left|D \ln \frac{d\widehat{\nu}}{d\widehat{\mu}}\right|_{H_S}^2\right]$$

Then we have

$$2J(\widehat{\nu}|\widehat{\mu}) = E_{\widehat{\nu}}\left[|\dot{v}_1|_{H_S}^2\right]$$

where v is the Schrödinger drift associated with ν. In particular we get the logarithmic-Sobolev inequality

$$H(\widehat{\nu}|\widehat{\mu}) \leq J(\widehat{\nu}|\widehat{\mu})$$

Proof: Let ν be the probability associated with $\widehat{\nu}$ by Definition IV.2 and v be the Girsanov drift of ν which is the free Schrödinger drift associated with $\widehat{\nu}$. Then Proposition V.1 and Proposition I.3 yield

$$H(\widehat{\nu}|\widehat{\mu}) = H(\nu|\mu) \leq J(\nu|\mu) \leq J(\widehat{\nu}|\widehat{\mu})$$

Furthermore Proposition V.1 and Proposition I.2 directly imply

$$d(\widehat{\mu},\widehat{\nu}) \leq d(\nu,\mu) \leq 2H(\nu|\mu) = 2H(\widehat{\nu}|\widehat{\mu})$$

□

2. Information loss on the path space

Let $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathbb{P})$ be a complete probability space, and let $(\mathcal{F}_t^{\mathbb{P}})_{t\in[0,1]}$ be a continuous filtration which satisfies the usual conditions. On this space let $(B_t)_{t\in[0,1]}$ be a $(\mathcal{F}_t^{\mathbb{P}})_{t\in[0,1]}$ and $S-$valued Brownian motion starting from the origin. In the point of view of information theory one may think of B as a Gaussian noise in a transmission channel. Assume now that one sends a signal u through this channel, and that the receptor observes $U = B + u$ but not u. Let $(\mathcal{F}_t^u)$ be the augmentation with respect to $\mathbb{P}$ of the filtration generated by the observed signal $t \to U_t$. If the receptor tries to estimate dynamically u, his best estimation will be

$$\widehat{u}_t = \int_0^t E_{\mathbb{P}}\left[\dot{u}_s | \mathcal{F}_s^u\right] ds$$

which is usually called the causal estimate of u. Hence the estimated signal will be $\widehat{u}$, while the emitted signal is u. Since $\widehat{u}$ is a projection of u, the energy of u is always bigger that the energy of $\widehat{u}$. With other words, some energy will dissipate in the channel, and the value of this dissipated energy is

$$E_{\mathbb{P}}\left[|u|_H^2\right] - E_{\mathbb{P}}\left[|\widehat{u}|_H^2\right]$$

As a matter of fact the value of this dissipated energy may be seen equivalently as an error or a loss of information. Indeed

$$E_{\mathbb{P}}\left[|u|_H^2\right] - E_{\mathbb{P}}\left[|\widehat{u}|_H^2\right] = E_{\mathbb{P}}\left[|u - \widehat{u}|_H^2\right]$$

is equal to the error of the causal estimate. The Theorem V.2 (which was first proved in **[27]** in the case of the Wiener space) states that the dissipated energy (or information) only depends on two parameters : the energy of the signal, and the law of the observed signal through its relative entropy. This loss of information relies on the fact that the observer only gets the information of the filtration $(\mathcal{F}_t^u)$ generated by $t \to U_t \in S$ which is smaller than $(\mathcal{F}_t^{\mathbb{P}})_{t\in[0,1]}$. Moreover one expects the equality to occur if and only if one can reconstruct the Brownian path until t from $\mathcal{F}_t^u$. This is exactly what shows the equality case in Theorem V.2. Before going further we have to set some notations. We note $L^2((\Omega, \mathcal{F}^{\mathbb{P}}, \mathbb{P}), H))$ or when there are no ambiguity on the underlying filtered space $L^2(\mathbb{P}, H)$ the set of the measurable mapping $u : \Omega \to H$, such that $E_{\mathbb{P}}\left[|u|_H^2\right] < \infty$ We also define $L_a^2(\mathbb{P}, H)$, the subset of the $u \in L^2(\mathbb{P}, H)$ such that the mapping $t \to \dot{u}_t$ is adapted to $(\mathcal{F}_t^{\mathbb{P}})_{t\in[0,1]}$. We recall that we defined $\mathcal{F}$ (resp. $(\mathcal{F}_t)$) as the completion of the sigma field $\mathcal{B}(W)$ with respect to μ (resp. the augmentation with respect to μ of the sigma field generated by the coordinate process $t \to W_t$).

Definition V.1. *Let $(B_t)_{t\in[0,1]}$ be a $S-$valued $(\mathcal{F}_t^{\mathbb{P}})_{t\in[0,1]}$ Brownian motion starting from the origin on a complete probability space $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathbb{P})$ with a continuous filtration $(\mathcal{F}_t^{\mathbb{P}})_{t\in[0,1]}$ which satisfies the usual conditions. Given a $u \in L_a^2(\mathbb{P}, H)$ such that $U\mathbb{P} \sim \mu$ where $U := I_W + u$, we define $(\mathcal{F}_t^u)$ as being the filtration $(\sigma(U_s, s \le t))_{t\in[0,1]}$ augmented with respect to $\mathbb{P}$. We also note $L_u^2(\mathbb{P}, H)$ the subset of the $\widetilde{u} \in L_a^2(\mathbb{P}, H)$ such that $t \to \dot{\widetilde{u}}_t$ is adapted to $(\mathcal{F}_t^u)$. Moreover $\widehat{u}$ will denote the projection of u on $L_u^2(\mathbb{P}, H)$ which is a closed subspace of $L_a^2(\mathbb{P}, H)$*

As matter of fact $\widehat{u}$ is the dual predicable projection of u (see **[7]**) on the augmentation with respect to $\mathbb{P}$ of the filtration generated by $t \to U_t = B_t + u_t$. We recall that physically $\widehat{u}$ may be seen as the causal estimator of a signal u and is written

$$\widehat{u}_t = \int_0^t E_{\mathbb{P}}\left[\dot{u}_s | \mathcal{F}_s^u\right] ds$$

We then have :

Lemma V.1. *Let $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathbb{P})$ be a complete probability space with a continuous filtration $(\mathcal{F}_t^{\mathbb{P}})_{t\in[0,1]}$ on it which satisfies the usual conditions. Let $(B_t)_{t\in[0,1]}$ be a $S-$valued $\mathcal{F}_t^P-$Brownian motion on $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathbb{P})$ and*

$U: t \in [0,1] \to U_t \in S$ be any $(\mathcal{F}_t)-$adapted, continuous process. Further assume that $U\mathbb{P} \sim \mu$. Then we have $\mathbb{P}-a.s.$

$$\widehat{u} + v \circ U = 0 \tag{2.44}$$

where v is the Girsanov drift associated with $U\mathbb{P}$.

Proof: For any $\theta \in L^2_a(U\mathbb{P}, H)$, since $t \to V_t$ is an abstract Wiener process on $(W, \mathcal{F}, U\mathbb{P})$. Hence we have

$$\begin{aligned}
E_{\mathbb{P}}\left[< \theta \circ U, u >_H\right] &= E_{\mathbb{P}}\left[\int_0^1 < \dot{\theta}_s \circ U, \dot{u}_s >_{H_S} ds\right] \\
&= E_{\mathbb{P}}\left[\int_0^1 \dot{\theta}_s \circ U dB_s + \int_0^1 < \dot{\theta}_s \circ U, \dot{u}_s >_{H_S} ds - \int_0^1 \dot{\theta}_s \circ U dB_s\right] \\
&= E_{\mathbb{P}}\left[(\int_0^1 \dot{\theta}_s dW_s) \circ U - \int_0^1 \dot{\theta}_s \circ U dB_s\right] \\
&= E_{\mathbb{P}}\left[(\int_0^1 \dot{\theta}_s dW_s) \circ U\right] \\
&= E_{U\mathbb{P}}\left[\int_0^1 \dot{\theta}_s dW_s\right] \\
&= E_{U\mathbb{P}}\left[\int_0^1 \dot{\theta}_s dW_s + \int_0^1 < \dot{\theta}_s, \dot{v}_s >_{H_S} ds - \int_0^1 < \dot{\theta}_s, \dot{v}_s >_{H_S} ds\right] \\
&= E_{U\mathbb{P}}\left[\int_0^1 \dot{\theta}_s dV_s - \int_0^1 < \dot{\theta}_s, \dot{v}_s >_{H_S} ds\right] \\
&= -E_{U\mathbb{P}}\left[\int_0^1 < \dot{\theta}_s, \dot{v}_s >_{H_S} ds\right] \\
&= -E_{\mathbb{P}}\left[< v \circ U, \theta \circ U >_H\right]
\end{aligned}$$

This shows (2.44) □

The following extends slightly the main result of **[46]**

Theorem V.2. *Let $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathbb{P})$ be a complete probability space with a continuous filtration $(\mathcal{F}^{\mathbb{P}}_t)_{t\in[0,1]}$ which satisfies the usual conditions. Let $(B_t)_{t\in[0,1]}$ be a $S-$valued $(\mathcal{F}^P_t)-$Brownian motion on $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathbb{P})$ and $U : [0,1] \to S$ be any $(\mathcal{F}^P_t)$-adapted, continuous process such that $U\mathbb{P} \sim \mu$. Then*

$$2H(U\mathbb{P}|B\mathbb{P}) \leq E_{\mathbb{P}}\left[|U - B|^2_H\right] \tag{2.45}$$

We further note $u := U - B$ and henceforth assume that $u \in L^2(\mathbb{P}, H)$. If we note $\widehat{u}$ the dual predicable projection of u on $(\mathcal{F}^u_t)$ which is the filtration $\sigma(U_s, s \leq t)$ augmented with respect to $\mathbb{P}$, we then have

$$2H(U\mathbb{P}|\mu) = E_{\mathbb{P}}\left[|U - B|^2_H\right] - \epsilon^2_U \tag{2.46}$$

where

$$\epsilon_U = \sqrt{E_{\mathbb{P}}\left[|u - \widehat{u}|^2_H\right]}$$

Morever the following assertions are equivalent

- $\epsilon_U = 0$
- *$V \circ U = B$ on $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathbb{P})$ where $V = I_W + v$ is the Girsanov shift associated with the measure $U\mathbb{P}$ on the Wiener space.*

Finally for any probability ν which is equivalent to the Wiener measure μ

$$2H(\nu|\mu) = \inf\left(\left\{E_{\mathbb{P}}\left[|U - B|^2_H\right]\right\}\right) \tag{2.47}$$

where the infimum is taken on all the $(\Omega, \mathcal{F}^{\mathbb{P}}, (\mathcal{F}_t^{\mathbb{P}})_{t\in[0,1]}, \mathbb{P})$ *and all the* (U, B) *defined on it as above and such that* $U\mathbb{P} = \nu$. *Moreover, we can always find (at least) one space* $(\Omega, \mathcal{F}^{\mathbb{P}}, (\mathcal{F}_t^{\mathbb{P}})_{t\in[0,1]}, \mathbb{P})$ *and a* (U, B) *defined on it with the same properties as in the first part of the claim, which attains the infimum.*

Proof: If $U - B \notin L^2(\mathbb{P}, H)$ then (2.45) holds. Otherwise let $u := U - I_W$ and let v be the Girsanov shift associated to $U\mathbb{P}$. By Proposition I.1 and Lemma V.1 we have :

$$\begin{aligned}
2H(U\mathbb{P}|\mu) &= E_{U\mathbb{P}}\left[|v|_H^2\right] \\
&= E_{\mathbb{P}}\left[|v\circ U|_H^2\right] \\
&= E_{\mathbb{P}}\left[|\widehat{u}|_H^2\right] \\
&= E_{\mathbb{P}}\left[|u|_H^2\right] - E_{\mathbb{P}}\left[|u-\widehat{u}|_H^2\right]
\end{aligned}$$

where the last line follows from the fact that $\widehat{u}$ is an orthogonal projection of u. This clearly yields (2.46) and the equivalences in case of equality. The inequality (2.45) clearly yields the inequality of (2.47). By taking $(\Omega, \mathcal{F}^{\mathbb{P}}, (\mathcal{F}_t^{\mathbb{P}})_{t\in[0,1]}, \mathbb{P}) = (W, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, \nu)$, and $(U, B) = (I_W, V)$, where $V = I_W + v$ is the Girsanov shift associated with ν, we know (see Proposition I.1) that the optimum is attained in (2.47). $\square$

Remark V.1. *In Theorem V.2 we wanted to stress that the physical origin of the inequality in (2.45) is the information loss. This is the reason why we need. Lemma V.1. We already showed this result in* **[27]** *(the underlying space was the Wiener space but the proof is the same). However the inequality (2.45) as well as the fact that the equality in (2.45) occurs if and only if* $V\circ U = B$ *may be showed directly. Indeed these results follow easily from Proposition I.1 and the Cauchy-Schwarz inequality as we showed it in* **[28]**. *Henceforth, let us assume that the underlying probability space is the Wiener space* W. *As we have seen, under the condition* $E_\mu\left[\rho(-\delta^W u)\right] = 1$, $U := I_W + u$ *is the inverse of* V *if and only if it is the right inverse of* V. *Hence the equality in (2.45) is equivalent to the invertibility and we recover the main result of* **[47]**. *For that reason we refer to this result as Üstünel's criterion.*

The Proposition V.3 is a path space version of the Brascamp-Lieb inequality. In Proposition V.4 we apply it to an h-path process, so that we get a Brascamp-Lieb inequality which holds on any abstract Wiener space. The proof of it is a generalization of the one given in **[26]**, however the ideas are essentially the same. As Theorem V.2 enlighten it the next inequalities are involved by information loss in the Gaussian channel.

Proposition V.3. *Let* $(\widetilde{\pi}^i)$ *be a family of projections on* H_S. *Further assume that for any* i, *there is a* $(e_j^i)_{j=1}^\infty$ *which is an Hilbert basis of* H_S *and a* $(I_i) \subset \mathbb{N}$ *such that* $\widetilde{\pi}_i = \sum_{j\in I_i} < x, e_j^i >_{H_S} e_j^i$ *and that*

$$\sum_i \alpha_i \widetilde{\pi}_i = I_{H_S}$$

for a sequence of positive numbers $(\alpha_i) \subset \mathbb{R}^+$. *We then set*

$$\widetilde{T}_i x = \sum_{j\in I_i} < e_j^i, x > e_j^i$$

which is well defined as a measurable mapping from $L^0(\widehat{\mu}, S) \to L^0(\widehat{\mu}, S)$. *We also note* $\pi_i : H \to H$ *the mapping such that* $(\pi_i h)(t) = \widetilde{\pi}_i(h_t)$ *for any* $(h, t) \in H \times [0, 1]$ *and* $T_i : \omega \in W \to T_i(\omega) \in W$ *the mapping defined pathwise by* $T_i(\omega) : t \in [0, 1] \to \widetilde{T}_i(W_t)$. *Then for any measure* ν *equivalent to* μ *with* $H(\nu|\mu) < \infty$ *we have*

$$H(\nu|\mu) \geq \sum_i \alpha_i H(T_i\nu|T_i\mu) \tag{2.48}$$

In particular for every $t \in [0, 1]$ *we have*

$$H(\nu|\mu) \geq \sum_i \alpha_i H(\widetilde{T}_i\nu_t|\widetilde{T}_i\mu_t) \tag{2.49}$$

where $\nu_t := W_t\nu$ (resp. $\mu_t := W_t\mu$) is the marginal at t.

Proof: From the definitions it is straightforward to check that (π^i) is a family of projections on H such that $\sum_i \alpha_i\pi^i = I_H$. Let (U, B) be the pair defined on a space $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathbb{P})$ which attains the optimum in the variational problem given in Theorem V.2, and set $u := U - B$. Theorem V.2 yields

$$\begin{aligned}
2H(\nu|\mu) &= E_{\mathbb{P}}\left[|u|_H^2\right] \\
&= E_{\mathbb{P}}\left[< \sum_i \alpha_i\pi_i u, u >_H\right] \\
&= \sum_i \alpha_i E_{\mathbb{P}}\left[|\pi_i u|_H^2\right] \\
&= \sum_i \alpha_i E_{\mathbb{P}}\left[|B + \pi_i u - B|_H^2\right] \\
&\geq \sum_i \alpha_i H((B + \pi_i u)\mathbb{P}|\mu) \\
&\geq \sum_i \alpha_i H(T_i(B + \pi_i u)\mathbb{P}|T_i\mu)
\end{aligned}$$

where the last equality comes from the fact that $H(\nu|\mu) \geq H(X\nu|X\mu)$ for any measurable $X : W \to E$ where E is a Polish space. By definition, we also have $T^i \circ \pi^i h = \pi^i h$ for any $h \in H$ so that $H(T_i(B+\pi_i u)\mathbb{P}|T_i\mu) = H(T_i U\mathbb{P}|T_i\mu) = H(T_i\nu|T_i\mu)$. Therefore we have (2.48). Since by definition for any $t \in [0, 1]$ $\mu - a.s.$ $W_t \circ T^i = \widetilde{T}_i(W_t)$, we also have

$$H(T_i\nu|T_i\mu) \geq H(W_t T_i\nu|W_t T_i\mu) = H(\widetilde{T}^i\nu_t|\widetilde{T}_i\mu_t)$$

□

Proposition V.4. *Let $(\widehat{\pi}^i)$ be a family of projections on H_S. Further assume that for any i there is a $(e_j^i)_{j=1}^{\infty}$ which is an Hilbert basis of H_S and a $(I_i) \subset \mathbb{N}$ such that $\widehat{\pi}_i = \sum_{j\in I_i} < x, e_j^i > e_j^i$, and that*

$$\sum_i \alpha_i\widehat{\pi}_i = I_{H_S}$$

for a sequence of positive numbers $(\alpha_i) \subset \mathbb{R}^+$ We define

$$\widehat{T}_i x = \sum_{j\in I_i} < e_j^i, x > e_j^i$$

which is well defined as a mapping from $L^0(\widehat{\mu}, S) \to L^0(\widehat{\mu}, S)$. Then for any measure $\widehat{\nu}$ equivalent to $\widehat{\mu}$ such that $H(\widehat{\nu}|\widehat{\mu}) < \infty$ we have

$$H(\widehat{\nu}|\widehat{\mu}) \geq \sum_i \alpha_i H(\widehat{T}_i\widehat{\nu}|\widehat{T}_i\widehat{\mu})$$

Proof: Let ν be the measure associated with $\widehat{\nu}$ by Definition IV.2 so that $W_1\nu = \widehat{\nu}$. By Proposition V.1 $H(\widehat{\nu}|\widehat{\mu}) = H(\nu|\mu)$. Hence Proposition V.3 with $t = 1$ implies

$$H(\widehat{\nu}|\widehat{\mu}) = H(\nu|\mu) \geq \sum_i \alpha_i H(\widehat{T}_i\widehat{\nu}|\widehat{T}_i\widehat{\mu})$$

□

Remark V.2. *Note that in the case $S = \mathbb{R}^d$ we have $\widehat{\pi}_i = \widehat{T}_i$ and if we note λ the Lebesgue measure and if we assume that $\widehat{\nu} = X\lambda$ for a X, we have $H(X\lambda|\widehat{\mu}) \geq \sum_i \alpha_i H((\widehat{\pi}^i X)\lambda|\widehat{\pi}^i\widehat{\mu})$ where $\widehat{\pi}^i\widehat{\mu}$ is the law of a standard Gaussian vector with range in $\widehat{\pi}(H_S)$. The relationship of this equality with the Brascamp-Lieb inequality was shown in* **[4]** *as it is also recalled in* **[26]**.

3. Üstünel's criterion in terms of variance and Shannon's inequality

In finite dimension, Shannon's inequality involves some entropies with respect of the Lebesgue measure. When we seek to write it in terms of Gaussian measures some correlation terms appear since we then lose the invariance under translations. Here we use a trick to recover the property of invariance under translation by performing a change of measure. Under this change of measure Theorem V.2 takes the form of Corollary V.1. The variational formulation of the entropy is then written in terms of variance instead of in terms of energy. In Theorem V.3 we get the abstract Wiener space version of the Shannon inequality as a consequence of two facts : the information loss on the path space, and the addition property of the variances of independent random variables.

Corollary V.1. *Let $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathbb{P})$ be a complete probability space with a continuous filtration $(\mathcal{F}_t^{\mathbb{P}})_{t\in[0,1]}$ which satisfies the usual conditions. Let $(B_t)_{t\in[0,1]}$ be a $\mathcal{F}_t^P$ $S-$valued Brownian motion on that space, and $U : [0,1] \to S$ be any $(\mathcal{F}_t)$-adapted continuous process such that $U\mathbb{P} \sim \mu$. Further assume that U is of the form $U := B + u$ where $u \in L_a^2(\mathbb{P}, H)$, and let μ_U be the probability defined by*

$$\frac{d\mu_U}{d\mu} := \rho(\delta^W m_U)$$

where $m_U(t) = \int_0^t E_{\mathbb{P}}[\dot{u}_s]\, ds$ Then

$$2H(U\mathbb{P}|\mu_U) \leq \mathcal{V}ar_{\mathbb{P}}(u)$$

where

$$\mathcal{V}ar_{\mathbb{P}}(u) = E_{\mathbb{P}}\left[|u - m_U|_H^2\right]$$

Let $\widehat{u}$ be the dual predicable projection of $u := U - B$ on $(\mathcal{F}_t^u)$ which is the filtration $\sigma(U_s, s \leq t)$ augmented with respect to $\mathbb{P}$. Then we have

$$2H(U\mathbb{P}|\mu_U) = \mathcal{V}ar_{\mathbb{P}}(u) - \epsilon_U^2$$

where

$$\epsilon_U = \left(E_{\mathbb{P}}\left[|u - \widehat{u}|_H^2\right]\right)^{\frac{1}{2}}$$

Moreover, the following assertions are equivalent

- $u = \widehat{u}$
- $V \circ U = B$ *where* $V = I_W + v$ *is the Girsanov shift associated with* $U\mathbb{P}$.
- $2H(U\mathbb{P}|\mu_U) = \mathcal{V}ar_{\mathbb{P}}(u)$
- $2H(U\mathbb{P}|\mu) = E_{\mathbb{P}}\left[|U - B|_H^2\right]$

Finally for any probability ν which is equivalent to the Wiener measure μ,

$$\frac{d\mu_\nu}{d\mu} := \rho(\delta^W m_\nu)$$

where $m_\nu(t) = -\int_0^t E_\nu[\dot{v}_s]\, ds = -E_\nu[v_t]$ Then

$$2H(\nu|\mu_\nu) = \inf\left(\{\mathcal{V}ar_{\mathbb{P}}(U - B)\}\right)$$

where the infimum is taken on all the $(\Omega, \mathcal{F}^{\mathbb{P}}, (\mathcal{F}_t^{\mathbb{P}})_{t\in[0,1]}, \mathbb{P})$ and all the (U, B) defined on it as above and such that $U\mathbb{P} = \nu$. Moreover, we can always find (at least) one space $(\Omega, \mathcal{F}^{\mathbb{P}}, (\mathcal{F}_t^{\mathbb{P}})_{t\in[0,1]}, \mathbb{P})$ and a (U, B) defined on it with the same properties as in the first part of the claim, which attains the infimum. Finally the optimum is attained by the same shifts as the variationnal problem of Theorem V.2

Proof: From the definitions, and Theorem V.2 we have

$$\begin{aligned} 2H(U\mathbb{P}|\mu_U) &= 2H(U\mathbb{P}|\mu) - 2E_{\mathbb{P}}\left[\ln \frac{d\mu_U}{d\mathbb{P}} \circ U\right] \\ &= 2H(U\mathbb{P}|\mu) - |m_U|_H^2 \\ &= E_{\mathbb{P}}\left[|\widehat{u}|_H^2\right] - |m_U|_H^2 \\ &= E_{\mathbb{P}}\left[|\widehat{u}|_H^2\right] - E_{\mathbb{P}}\left[|u|_H^2\right] - \left(E_{\mathbb{P}}\left[|u|_H^2\right] - |m_U|_H^2\right) \end{aligned}$$

which is the main part of the result. Note that if $U\mathbb{P} = \nu$ we have

$$\begin{aligned} m_U &= \int_0^{\cdot} E_{\mathbb{P}}\left[\dot{u}_s\right] ds \\ &= \int_0^{\cdot} E_{\mathbb{P}}\left[\dot{\widehat{u}}_s\right] ds \\ &= -\int_0^{\cdot} E_{\mathbb{P}}\left[\dot{v}_s \circ U\right] ds \\ &= -\int_0^{\cdot} E_{\nu}\left[\dot{v}_s \circ U\right] ds \\ &= m_\nu \end{aligned}$$

so that $\mu_\nu = \mu_U$. By taking $(\Omega, \mathcal{F}^{\mathbb{P}}, (\mathcal{F}_t^{\mathbb{P}})_{t\in[0,1]}, \mathbb{P}) = (W, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, \nu)$, and $(U, B) = (I_W, V)$, where $V = I_W + v$ is the Girsanov shift associated with ν, we know (see Proposition I.1) that the optimum is attained □

The next lemma is the price to pay for working with Gaussian measures instead of the Lebesgue measure (which is not defined when $\dim S = \infty$)

Lemma V.2. *Let ν be a probability equivalent to μ such that $H(\nu|\mu) < \infty$, and let $\widehat{\nu}$ a probability equivalent to $\widehat{\mu}$. Further assume that*

$$W_1\nu = \widehat{\nu}$$

and let

$$\frac{d\mu_\nu}{d\mu} := \rho(\delta^W m_\nu)$$

where $m_\nu(t) = -\int_0^t E_\nu\left[\dot{v}_s\right] ds = -E_\nu\left[v_t\right]$ where v is the Girsanov drift associated with ν, $\nu_t = W_t\nu$ and

$$\frac{d\widehat{\mu}_{\widehat{\nu}}}{d\widehat{\mu}}(x) := \exp\left(< x, m_{\widehat{\nu}} > - \frac{|m_{\widehat{\nu}}|_{H_S}^2}{2}\right)$$

where $m_{\widehat{\nu}}$ is the mean of $\widehat{\nu}_i$. Then

$$H(\widehat{\nu}|\widehat{\mu}_{\widehat{\nu}}) \leq H(\nu|\mu_\nu) \tag{3.50}$$

Moreover, if ν is the measure associated with $\widehat{\nu}$ by Definition IV.2 the inequality (3.50) is an equality.

Proof: Since $\widehat{\mu}$ is centered ,$V_1\nu = \widehat{\mu}$, and $W_1\nu = \widehat{\nu}$ we have

$$m_\nu(1) = -E_\nu\left[v_1\right] = \int d\widehat{\nu}(x) x = m_{\widehat{\nu}}$$

Hence by applying two times the Cameron-Martin theorem (see **[25]**) once on W, once on S we obtain

$$W_1\mu_\nu = (W_1 + m_\nu(1))\mu = (W_1 + m_{\widehat{\nu}})\mu = (I_S + m_{\widehat{\nu}})\widehat{\mu} = \widehat{\mu}_{\widehat{\nu}}$$

Since we also have $W_1\nu = \widehat{\nu}$, we then obtain

$$H(\widehat{\nu}|\widehat{\mu}_{\widehat{\nu}}) = H(W_1\nu|W_1\mu_\nu) \leq H(\nu|\mu_\nu)$$

Moreover if ν is the measure associated with $\widehat{\nu}$ by Definition IV.2 it is straightforward to check that we have an equality. □

We now give an abstract Wiener space version of Shannon's inequality.

Theorem V.3. *Let $(\widehat{\nu}^i)$ be a sequence of probabilities equivalent to $\widehat{\mu}$ such that $H(\widehat{\nu}^i|\widehat{\mu}) < \infty$. For a sequence (p_i) of positive reals such that*

$$\sum_i p_i = 1$$

we set $\widehat{\nu}^\Sigma := (\sum_i \sqrt{p_i}\pi_i) \otimes_i \widehat{\nu}^i$ where π_i is the projection on the i-th coordinate of the product space $S^{\mathbb{N}}$. We further define a family of measure $(\widehat{\mu}^i)$ by

$$\frac{d\widehat{\mu}^i}{d\widehat{\mu}} = \exp\left(< x, \widehat{m}_i > - \frac{|\widehat{m}_i|^2_{H_S}}{2} \right)$$

where $\widehat{m}_i$ is the mean of $\widehat{\nu}^i$ and we set

$$\frac{d\widehat{\mu}^\Sigma}{d\widehat{\mu}} = \exp\left(< x, \sum_i \sqrt{p_i}\widehat{m}_i > - \frac{|\sum_i \sqrt{p_i}\widehat{m}_i|^2_{H_S}}{2} \right)$$

Then we have

$$H(\widehat{\nu}^\Sigma|\widehat{\mu}^\Sigma) \leq \sum_i p_i H(\widehat{\nu}^i|\widehat{\mu}^i)$$

Proof: For any $i \in \mathbb{N}$ let ν^i be the optimal measure associated with $\widehat{\nu}^i$ by Definition IV.2, and let V^i be the Girsanov shift associated with ν^i. For any i we also note μ^i the measure associated with ν^i on W by Lemma V.2. From Corollary V.1 we have

$$H(\nu^i|\mu^i) = \mathcal{V}ar_{\nu^i}\left[(V^i - I_W) \right]$$

Thus Lemma V.2 yields

$$H(\widehat{\nu}^i|\widehat{\mu}^i) = \mathcal{V}ar_{\nu^i}\left[(V^i - I_W) \right] \tag{3.51}$$

We set $\Omega^\Sigma = W^{\mathbb{N}}$, $\mathbb{P}^\Sigma := \otimes_i \nu^i$, and we define the filtration $(\mathcal{G}^\Sigma_t)$ by

$$\mathcal{G}^\Sigma_t = \sigma\left(\left\{ V^i_s \circ Pr_i \right\}_{i\in\mathbb{N}}, s \leq t \right\} \right)$$

for any $t \in [0,1]$, where Pr_i is the projection on the $i-th$ coordinate of $W^{\mathbb{N}}$. We also set $\mathcal{G}^\Sigma := \mathcal{G}^\Sigma_1$. Then from Paul Levy's theorem

$$B^\Sigma := \sum_i \sqrt{p_i} V^i \circ Pr_i$$

is a $(\mathcal{G}^\Sigma_t)-$Brownian motion on $(\Omega^\Sigma, \mathcal{G}^\Sigma, \mathbb{P}^\Sigma)$. On the other hand $U^\Sigma := \sum_i \sqrt{p_i} Pr_i$ is adapted to $(\mathcal{G}^\Sigma_t)$. Hence Corollary V.1 applies and we get

$$H(U^\Sigma \mathbb{P}^\Sigma|\mu_\Sigma) \leq \mathcal{V}ar_{\mathbb{P}^\Sigma}(U^\Sigma - B^\Sigma) \tag{3.52}$$

Where μ^Σ is the measure defined in Corollary V.1. Since $W_1(U^\Sigma\mathbb{P}^\Sigma) = W_1 \sum_i \sqrt{p_i} Pr_i \mathbb{P}^\Sigma = \widehat{\nu}^\Sigma$ Lemma V.2 yields

$$H(\widehat{\nu}^\Sigma|\widehat{\mu}_\Sigma) \leq \mathcal{V}ar_{\mathbb{P}^\Sigma}(U^\Sigma - B^\Sigma) \tag{3.53}$$

The property of the variance of a sum of independant variables writes

$$\mathcal{V}ar_{\mathbb{P}^\Sigma}(U^\Sigma - B^\Sigma) = \sum_i p_i \mathcal{V}ar_{\nu^i}(V^i - I_W) \tag{3.54}$$

By gathering equations (3.51), (3.53) and (3.54) we get the result. □

Remark V.3. *In the finite dimensional case $S = \mathbb{R}^n$ with the Lebesgue measure λ on it, let (X_i) be a sequence of independent random elements with values in $\mathbb{R}^n$ defined on a space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $H(X_i\mathbb{P}|\lambda) < \infty$. Shannon's inequality can be written*

$$\sum_i p_i H(X^i\mathbb{P}|\lambda) \geq H(\sum_i \sqrt{p_i} X_i\mathbb{P}|\lambda) \tag{3.55}$$

where (p_i) is a sequence of positive numbers such that $\sum_i p_i = 1$. However, the Lebesgue measure is not defined in infinite dimensions and we had to write it in terms of Gaussian measure (which still makes sense in infinite dimension as a Wiener measure). Let γ be the standard Gaussian measure on $\mathbb{R}^n$. The trick to keep a formula as simple as possible is then to introduce the following measures : for any i we set

$$\frac{d\gamma^i}{d\lambda}(x) := \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{(x - E\left[X_i\right])^2}{2}\right)$$

and

$$\frac{d\gamma^\Sigma}{d\lambda}(x) := \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{(x - \sum_i \sqrt{p_i} E\left[X_i\right])^2}{2}\right)$$

We then have

$$\begin{aligned} H(X^i\mathbb{P}|\gamma^i) &= E_{X^i\mathbb{P}}\left[\ln \frac{dX^i\mathbb{P}}{d\lambda}\right] - E_{X^i\mathbb{P}}\left[\ln \frac{d\gamma^i}{d\lambda}\right] \\ &= H(X^i\mathbb{P}|\lambda) - E_{X^i\mathbb{P}}\left[\ln \frac{d\gamma^i}{d\lambda}\right] \\ &= H(X^i\mathbb{P}|\lambda) - \int d\lambda(x) \frac{(X^i(x) - E\left[X^i\right])^2}{2} - \frac{n}{2}\ln(2\pi) \\ &= H(X^i\mathbb{P}|\lambda) - \mathcal{V}ar_{\mathbb{P}}(X^i) - \frac{n}{2}\ln(2\pi) \end{aligned}$$

Since the X^i are independent we have

$$\mathcal{V}ar_{\mathbb{P}}(\sum_i \sqrt{p_i} X^i) = \sum_i p_i \mathcal{V}ar_{\mathbb{P}}(X_i)$$

Hence if we set $\nu^\Sigma = \sum_i \sqrt{p_i} X_i \mathbb{P}$ we get

$$\sum_i p_i H(X^i\mathbb{P}|\gamma^i) - H(\nu^\Sigma|\gamma^\Sigma) = \sum_i p_i H(X^i\mathbb{P}|\lambda) - H(\nu^\Sigma|\lambda)$$

Hence Shannon's inequality may be written $H(\nu^\Sigma|\gamma^\Sigma) \leq \sum_i p_i H(X^i\mathbb{P}|\gamma^i)$. This is exactly what we proved in this proposition.

Part 2

A notion of stochastic invertibility on Wiener space

VI

Invertibility of Girsanov shifts on abstract Wiener space : general case

Abstract. **Introduction. Stochastic calculus on abstract Wiener space with an abstract time structure. The innovation process. $\nu-$invertibility of adapted shifts. $\nu-$invertibility and entropy. Local properties. Connexion with optimal transport. Connexion with the innovation conjecture.**

1. Introduction

In the first part of this work we considered the invertibility of a shift $V := I_W + v$, where v is an adapted drift on a classical Wiener space (W, H, μ) which satisfies the two hypothesis of **[53]**, namely

$$v \in L_a^0(\mu, H)$$

and

$$E_\mu\left[\rho\left(-\delta^W v\right)\right] = 1$$

Under these two hypothesis the notion of invertibility was well defined by using μ almost surely defined elements, and it was proved to be useful in various applications in Part I. In particular, we recalled that this notion provides naturally variational formulas for the relative entropy $H(\nu|\mu)$ in the case $\nu \sim \mu$. A natural question is whether one can extend these formulas to the case of a probability ν which is absolutely continuous, but not equivalent to the Wiener measure μ. We were interested in the invertibility of any $V := I_W + v$, where v satisfies the two above hypothesis. In the case of a measure $\nu << \mu$ we will see that we are not interested in the invertibility of any $V := I_W + v$, $v \in L_a^0(\nu, H)$, but that we have to focus on the invertibility of a particular shift associated with ν : the Girsanov shift. Indeed when we considered a measure ν equivalent to μ it appeared that the relevant associated shift v was such that $\mu-$a.s.

$$\frac{d\nu}{d\mu} = \rho\left(-\delta^W v\right) \tag{1.56}$$

In the case of the classical Wiener space, it is well known that this result also holds in the case of a measure ν absolutely continuous with respect to μ but not necessarily equivalent. More accurately, it is well known that for any $\nu << \mu$ we can find a $v \in L_a^0(\nu, H)$ such that $\nu - a.s.$ (1.56) holds. We still call $V := I_W + v$ the Girsanov shift associated with ν. Under the scope of the stochastic invertibility we introduced in the introduction to the Thesis, we recall that V is what we called the Brownian transform. However in the case of probabilities absolutely continuous with respect to the Wiener measure, the name Girsanov shift seems to be more suggestive. As a matter of fact we will work on a more abstract structure : an abstract Wiener space where the time structure is given by a continuously increasing sequence of projections on the Cameron-Martin space. In that case, the existence of a Girsanov shift for any measure absolutely continuous with respect to the Wiener measure becomes non trivial. However, the poor structure of that space has the advantage to force us to develop new proofs which appear most of the time to be more straightforward, and more fundamentals than the known ones : we used the spirit of these proofs in Part I. Within this context, we will see that we can still define a notion of invertibility for the Girsanov shifts of probabilities absolutely continuous with respect to the Wiener measure, and we will extend the study of invertibility within this context : some other relations between invertibility, optimal transport, and information theory will also be considered here. Furthermore, we also study of the local properties in that case. Here, we emphasize that these general theorems enable to extend the results of part I to probabilities absolutely continuous with respect to the Wiener measure. The main reason why we did not investigate the applications directly in the case of measures absolutely continuous is that we did not want to disturb a potential reader only interested in the applications, with some technical difficulties typically related to the fact that the negligible sets of the probability may be larger than those of the Wiener measure.

The structure of the present chapter is the following. In Section 2 we recall some notions of abstract stochastic calculus which are presented with much details in [**50**], and we also add some easy but useful results. We also recall a result of optimal transport on Wiener space which will be used in the sequel. In Section 3 we prove the existence of a Girsanov shift in this abstract framework for probabilities absolutely continuous with respect to the Wiener measure. Although part of the claim already appeared in [**50**], the proof was wrong in the general case. In Section 4 we study the pull-backs of the Girsanov shifts with some related morphisms of probability space. The interesting point is our proofs which are very short, while the results are very general. We want to emphasize that as long as we consider the stochastic processes as some morphisms of probability space, most of the connexions of invertibility with the applications (for instance, those investigated in [**47**]), can be obtained easily from these results. In Section 5, we give some easy applications of these results, we define the innovation process, and we recall its main properties. In Section 6, we investigate the connexion between energy, entropy, and mean square error. In Section 7 we relate this problem to the invertibility of Girsanov shifts, and we state a generalization of the main result of [**47**]. As a matter of fact, even on the classical case, our results are slightly more generals than the known versions. In Section 8 we extend the general results of part I related to the local properties of invertibility. In Section 9 and Section 10 we relate the invertibility to the Monge problem in a new way, and to the innovation conjecture. Both of these applications come from a polar decomposition which can be obtained if the Girsanov shift is invertible.

2. Preliminaries and notation

2.1. Abstract Wiener Space.

Let H be a separable Hilbert space and let W be the completion of H with respect to a measurable semi norm (see [**25**]). Then W is a separable Banach space and the injection i $H \hookrightarrow W$ is dense and continuous. By identifying H with its dual thanks to the Riesz representation, we note $i^\star$ the injection $W^\star \hookrightarrow H$, which is also dense and continuous.

$$W^\star \hookrightarrow^{i*} H \hookrightarrow^{i} W$$

In the sequel we shall always identify $W^\star$ with a subset of H by means of this latter injection and the notation $|l|_H$ will always mean $|i^\star(l)|_H$. Moreover $< .,. >_{W^\star,W}$ will denote the duality bracket between W and $W^\star$. The triplet (W,H,i) is an abstract Wiener space and H is the associated Cameron-Martin space. Let $\mathcal{B}(W)$ be the Borelian sigma field on W. It is well know that there is a unique Borelian probability μ on $(W,\mathcal{B}(W))$ such that for each $l \in W^\star$

$$\int \mu(dw) \exp\left(i < l, w >_{W^\star,W}\right) = \exp\left(-\frac{|l|_H^2}{2}\right)$$

This probability is called the Wiener measure. In particular, this means that every $l \in W^\star$ is in every $\mathcal{L}^p(\mu)$, $p \geq 1$. By noting $L^p(\mu)$ the equivalence classes which identify the elements of $\mathcal{L}^p(\mu)$ which are equal $\mu - as$, we can define $\delta : W^\star \subset H \to L^p(\mu)$ as being the mapping which associates to any $l \in W^\star$ its equivalence class in $L^p(\mu)$. By construction the Wiener measure is characterized by the fact that δl is a gaussian centered normal with a variance $|l|_H^2$, and the covariance between δl and δj is $< l, j >_H$. In other words $\delta : W^\star \subset H \to L^p(\mu)$ is an isometry. Now, recall that $W^\star$ may be viewed as a dense subspace of H. Hence δ extends as an isometry to an isometry on every $L^p(\mu)$, $p \geq 1$, which we also note $\delta : H \to L^p(\mu)$. More explicitly for each $h \in H$ there is a sequence $(l_n)_{n\in\mathbb{N}} \subset W^\star$ converging strongly to h in H. Therefore $(< l_n, \omega >_{W^\star, W})_{n\in\mathbb{N}} \subset L^p(\mu)$ is Cauchy in the complete space $L^p(\mu)$ for every $p > 1$, and hence converges in $L^p(\mu)$ to a random variable $\delta h = \lim_{n\to\infty} < l_n, \omega >_{W^\star, W}$ where the limit is in $L^p(\mu)$. In particular $\delta h(\omega + k) = \delta h + < h, k >_H$ for every $k, h \in H$. The set $\{\delta h, h \in H\}$ is an isonormal Gaussian field.

2.2. Stochastic calculus in Abstract Wiener Space.
In this section we recall the construction of a stochastic calculus on the abstract Wiener space by means of a sequence of continuous increasing projections which represents the arrow of time (for much details see [**50**]). Let $(\pi_t)_{t\in[0,1]}$ be a continuous and increasing sequence of orthogonal projections on H such that $\pi_0 = 0_H$, $\pi_1 = I_H$ and for any $t \in [0,1]$ $\pi_t(W^\star) \subset W^\star$. In this chapter we only consider Wiener spaces on which such sequences exist. An example of such spaces is provided by the classical Wiener space $C([0,1], \mathbb{R})$ with the family π_t defined by $\pi_t h := \int_0^{\cdot\wedge t} \dot{h}_s ds$ for any $h \in H$. We can also note that such structures also apply to the Gaussian Free Field which may be seen as an abstract Wiener space. In that case the sequence (π_t) may be given by any resolution of the identity $\pi : \mathcal{B}([0,1]) \to \mathcal{L}(H)$ ($\mathcal{L}(H)$ denotes the set of the linear operators on H) by taking $\pi_t := \pi([0,t])$ for any $t \in [0,1]$, whenever $(\pi([0,t]))_{t\in[0,1]}$ satisfies the above properties. For further interesting examples we refer to [**49**]. We further note $\mathcal{B}(W)$ be the Borelian $\sigma-$ field on W. Since we shall deal with a notion of "adapted" process, we will also have to consider $\mathcal{F}^\mu$ which is the completion of $\mathcal{B}(W)$ with respect to the Wiener measure μ. In order to handle $\mathcal{F}^\mu$, we note Θ^μ the $\mu-$negligible sets of $\mathcal{B}(W)$. We will also consider the measure $\overline{\mu}$ which is the extension of μ on $\mathcal{F}^\mu$. Let ν be another probability on $(W, \mathcal{B}(W))$ such that $\nu << \mu$ (i.e. absolutely continuous) we recall that ν has a unique extension $\overline{\nu}$ on $\mathcal{F}^\mu$ such that $\overline{\nu} << \overline{\mu}$. Moreover, ν has also an extension $\widetilde{\nu}$ on the completion $\mathcal{F}^\nu$ of $\mathcal{B}(W)$ with respect to ν. Furthermore Θ^ν will denote the $\nu-$negligible sets of $\mathcal{B}(W)$, and we have $\Theta^\mu \subset \Theta^\nu$ so that $\mathcal{F}^\mu \subset \mathcal{F}^\nu$. In the sequel $\overline{\mu}$ will be also noted μ, and will be called the Wiener measure : then it is defined on $(W, \mathcal{F}^\mu)$. When we shall introduce a measure ν such that $\nu << \mu$, ν will denote $\overline{\nu}$ i.e. it is defined on $(W, \mathcal{F}^\mu)$. Moreover we will also sometimes consider ν as a measure on $(W, \mathcal{F}^\nu)$, and in that case the letter ν will also represent what should be noted $\widetilde{\nu}$. If ν and $\widehat{\nu}$ are two probabilities which are absolutely with respect to the Wiener measure μ we note $\mathcal{M}\left(\left(W, \mathcal{F}^{\widehat{\nu}}\right), (W, \mathcal{F}^\nu)\right)$ the set of the measurable maps from $\left(W, \mathcal{F}^{\widehat{\nu}}\right)$ into $(W, \mathcal{F}^\nu)$, and $M_{\widehat{\nu}}\left(\left(W, \mathcal{F}^{\widehat{\nu}}\right), (W, \mathcal{F}^\nu)\right)$ the equivalence classes of elements of $\mathcal{M}\left(\left(W, \mathcal{F}^{\widehat{\nu}}\right), (W, \mathcal{F}^\nu)\right)$ which is constructed by identifying the elements which are equals $\widehat{\nu}-$ almost surely. In particular we note $\mathcal{L}^0(\mu, W) = \mathcal{M}((W, \mathcal{F}^\mu), (W, \mathcal{F}^\mu))$ (resp. for a $\nu << \mu$ we note $\mathcal{L}^0(\nu, W) = \mathcal{M}((W, \mathcal{F}^\nu), (W, \mathcal{F}^\mu)))$ and $L^0(\mu, W) = M_\mu((W, \mathcal{F}^\mu), (W, \mathcal{F}^\mu))$ (resp. $L^0(\nu, W) = M_\nu((W, \mathcal{F}^\nu), (W, \mathcal{F}^\mu)))$. We also note $L^0(\mu, H) = \left\{u \in L^0(\mu, W) : |u|_H < \infty \ \mu - as\right\}$ (resp. $L^0(\nu, H) = \left\{u \in L^0(\nu, W) : |u|_H < \infty \ \nu - as\right\}$) and $\mathcal{L}^0(\mu, H) = \left\{u \in \mathcal{L}^0(\mu, W) : |u|_H < \infty \ \mu - as\right\}$ (resp. $\mathcal{L}^0(\nu, H) = \left\{u \in \mathcal{L}^0(\nu, W) : |u|_H < \infty \ \nu - as\right\}$. $L^p(\nu, H)$ denote the subset of the $u \in L^0(\nu, H)$ such that $E_\nu\left[|u|_H^p\right] < \infty$. More generally we will often deal with some filtered probability space of the form $\left(W, \mathcal{F}, (\mathcal{G}_t)_{t\in[0,1]}, \nu\right)$ where $\mathcal{F}$ is a sigma field, $(\mathcal{G}_t)_{t\in[0,1]}$ a filtration on it, and ν is a probability on $(W, \mathcal{F})$. Whenever $\mathcal{G}_1 = \mathcal{F}$, we will note $\left(W, (\mathcal{G}_t)_{t\in[0,1]}, \nu\right)$ or $(W, \mathcal{G}_., \nu)$ instead of $\left(W, \mathcal{F}, (\mathcal{G}_t)_{t\in[0,1]}, \nu\right)$, and more generally $\mathcal{G}_.$ will denote the filtration $(\mathcal{G}_t)_{t\in[0,1]}$. Moreover, for any $t \in [0,1]$, $L^2(\mathcal{G}_t, \nu)$ will denote the equivalence classes of the set $\mathcal{L}^2(\mathcal{G}_t, \nu)$ of the $\mathcal{G}_t$ measurable functions which are square integrable with respect to ν, which we get by identifying the elements of $\mathcal{L}^2(\mathcal{G}_t, \nu)$ which are equal $\nu - a.s.$ Similarly we will deal, for a given $t \in [0,1]$, with the spaces $L^0(\mathcal{G}_t, \nu, H)$ $(L^2(\mathcal{G}_t, \nu, H))$ which are the equivalence class with respect to ν of the $\mathcal{G}_t$ measurable functions u such that $\nu - a.s.$ $u \in H$ (resp. such that $E_\nu\left[|u|_H^2\right] < \infty$). We denote by I_W the identity map

$$\begin{array}{cccc} I_W & : & W & \to & W \\ & & \omega & \mapsto & \omega \end{array}$$

Obviously $I_W \in L^0(\mu, W)$ and for every probability ν such that $\nu << \mu$ we also have $I_W \in L^0(\nu, W)$. Since W is separable, it is well known that its Borelian sigma field is also the sigma field generated by the weak open subsets of W, which is also the sigma field generated by the elements of $W^\star$. Hence, we can use I_W to build a canonical filtration.

Definition VI.1. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure (possibly $\nu = \mu$) and $X \in \mathcal{L}^0(\nu, W)$. The natural filtration generated by X with respect to the time sequence $\pi_.$, denoted by $\mathcal{F}_.^{X,0}$ is defined by*

$$\mathcal{F}_t^{X,0} = \sigma\left(< \pi_t l, X >,\ l \in W^\star\right)$$

for every $t \in [0,1]$. The associated augmented filtration $\mathcal{F}_.^X$ is defined by

$$\mathcal{F}_t^X = \sigma\left(\mathcal{F}_t^{X,0} \cup \Theta^\nu\right)$$

for every $t \in [0,1]$. Since we will often deal with it, we adopt a special notation for $\mathcal{F}_.^{I_W}$ (resp. $\mathcal{F}_.^{I_W,0}$) and we note it $\mathcal{F}_.^\nu$ (resp. $\mathcal{F}_.^{W,0}$). Note that $\mathcal{F}_1^\nu = \mathcal{F}^\nu$. Conversely, a random mapping $X \in L^0(\nu, W)$ will be said adapted to a filtration $\mathcal{G}_.$ if for every $t \in [0,1]$ $\mathcal{F}_t^X \subset \mathcal{G}_t$. If X is adapted to the particular filtration $\mathcal{F}_.^\nu$, it will be said to be $\nu-$adapted.

Note that contrary to the natural filtration, the augmented filtration only depends on the equivalence class of $X \in L^0(\nu, W)$ so that it is meaningfull to speak of the augmented filtration of any $X \in L^0(\nu, W)$. Furthermore, when $\mathcal{G}_.$ is a filtration on $\mathcal{F}^\nu$ where ν is a probability such that $\nu << \mu$, we note $L_a^0(\mathcal{G}_., \nu, H)$ (resp. $L_a^2(\mathcal{G}_., \nu, H)$) the subset of the u in $L^0(\nu, H)$ (resp. $L^2(\nu, H)$) such that u is adapted to $(\mathcal{G}_t)_{t\in[0,1]}$. Since we will often deal with it, we adopt a special notation for $L_a^0(\mathcal{F}_.^\nu, \nu, H)$ (resp. $L_a^2(\mathcal{F}_.^\nu, \nu, H)$) and we note it $L_a^0(\nu, H)$ (resp. $L_a^2(\nu, H)$).

Definition VI.2. *Let $\nu << \mu$ be a probability which is absolutely continuous with respect to the Wiener measure, $\mathcal{G}_.$ be a filtration on $(W, \mathcal{F}^\nu)$, and $M \in L^0(\nu, W)$. Then the triple $(M,\ H,\ \pi_.)$ is called an abstract martingale on $(W,\ \mathcal{G}_.,\ \nu)$ if for every l in $W^\star$ the process*

$$(t, \omega) \to \delta^M(\pi_t l)(\omega) := < \pi_t l, M(\omega) >_{W^\star, W}$$

is a martingale on $(W, \mathcal{G}_., \nu)$.

The next definition of an abstract Wiener process is based on the idea of **[32]**. As a matter of fact, it uses the Levy criterion to generalize the notion of Brownian motion :

Definition VI.3. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ, and $X \in L^0(\nu, W)$, and $\mathcal{G}_.$ be a filtration on $(W, \mathcal{F}^\nu)$. Then the triple $(X,\ H,\ \pi_.)$ is called an abstract Wiener process on $(W,\ \mathcal{G}_.,\ \nu)$ if the two following conditions hold :*

1. *$(X,\ H,\ \pi_.)$ is an abstract martingale on $(W,\ \mathcal{G}_.,\ \nu)$*
2. *$< \delta^M \pi_. l >_t = |\pi_t l|_H^2$*

Remark VI.1. *LetM be as in the last definition, then we also have $M\nu = \mu$. Indeed, assume that for every $l \in W^\star$, we have $\left\langle \delta^M \pi_. l \right\rangle_t = |\pi_t l|_H^2$. Then we also have :*

$$E_\nu\left[\exp\left(\frac{< \delta^M \pi_. l >_1}{2}\right)\right] = \exp\left(\frac{|l|_H^2}{2}\right) < \infty$$

Hence by the Novikov criterion $t \to \exp\left(i\delta^M \pi_t l - \frac{|i\pi_t l|_H^2}{2}\right)$ is a martingale on $(W, \mathcal{G}_., \nu)$. In particular for every $0 \leq s < t \leq 1$ we have

$$E_\nu\left[\exp\left(i\delta^M \pi_t l - \frac{|i\pi_t l|_H^2}{2}\right)\middle|\mathcal{G}_s\right] = \exp\left(i\delta^M \pi_s l - \frac{|i\pi_s l|_H^2}{2}\right)$$

Since $\pi_0 = 0$ and $\pi_1 = I_H$, the last equation with $s = 0$ and $t = 1$ gives :

$$E_\nu \left[\exp\left(i < l, M >_{W^\star, W}\right)\right] = \exp\left(-\frac{|l|_H^2}{2}\right)$$

which implies $M\nu = \mu$. Conversely, any abstract martingale $M \in L^0(\nu, W)$ on $(W, \mathcal{G}_., \nu)$ which is such that $M\nu = \mu$ is an abstract Wiener process with respect to its own filtration $\mathcal{F}_.^M$. Indeed we have for every $0 \leq s \leq t \leq 1$

$$\begin{aligned} E_\nu\left[\left(\delta^M \pi_t l\right)^2 \middle| \mathcal{F}_s^M\right] &= E_\nu\left[\left(\delta^M\left(\left(\pi_t - \pi_s\right)l\right)\right)^2 \middle| \mathcal{F}_s^M\right] + \left(\delta^M \pi_s l\right)^2 + 2\delta^M \pi_s l E_\nu\left[\delta^M\left(\left(\pi_t - \pi_s\right)l\right) \middle| \mathcal{F}_s^M\right] \\ &= E_\nu\left[\left(\delta^M\left(\left(\pi_t - \pi_s\right)l\right)\right)^2 \middle| \mathcal{F}_s^M\right] + \left(\delta^M \pi_s l\right)^2 + 2\delta^M \pi_s l E_\nu\left[E_\nu\left[\delta^M\left(\left(\pi_t - \pi_s\right)l\right) \middle| \mathcal{G}_s\right] \middle| \mathcal{F}_s^M\right] \\ &= E_\nu\left[\left(\delta^M\left(\left(\pi_t - \pi_s\right)l\right)\right)^2 \middle| \mathcal{F}_s^M\right] + \left(\delta^M \pi_s l\right)^2 \end{aligned}$$

On the other hand, for every $l, \widetilde{l} \in W^\star$ and $s < t$,

$$\begin{aligned} E_\nu\left[\delta^M\left(\left(\pi_t - \pi_s\right)l\right)\delta^M\left(\pi_s \widetilde{l}\right)\right] &= E_{M\nu}\left[\delta^W\left(\left(\pi_t - \pi_s\right)l\right)\delta^W\left(\pi_s \widetilde{l}\right)\right] \\ &= E_\mu\left[\delta^W\left(\left(\pi_t - \pi_s\right)l\right)\delta^W\left(\pi_s \widetilde{l}\right)\right] \\ &= < \left(\pi_t - \pi_s\right)l, \pi_s \widetilde{l} >_H \\ &= 0 \end{aligned}$$

Since for every $l \in W^\star$ the law of $\delta^M l$ on $(W, \mathcal{F}^\nu, \nu)$ is $\mathcal{N}\left(0, |l|_H^2\right)$, this implies that $\delta^M\left(\left(\pi_t - \pi_s\right)l\right)$ is independent of $\mathcal{F}_s^M$ with law $\mathcal{N}\left(0, |\left(\pi_t - \pi_s\right)l|_H^2\right)$. Finally we have :

$$\begin{aligned} E_\nu\left[\left(\delta^M \pi_t l\right)^2 - |\pi_t l|_H^2 \middle| \mathcal{F}_s^M\right] &= |\left(\pi_t - \pi_s\right)l|_H^2 + \left(\delta^M \pi_s l\right)^2 - |\pi_t l|_H^2 \\ &= \left(\delta^M \pi_s l\right)^2 - |\pi_s l|_H^2 \end{aligned}$$

Where the last equation comes from the equality $< \pi_s l, \pi_t l >_H = |\pi_s l|_H^2$, which holds for any $0 \leq s \leq t \leq 1$.

We recall the construction of the abstract stochastic integral over an abstract Wiener process. Since most of the time we will deal with it, and to avoid to heavy notations, we build the stochastic integral with respect to the canonical filtration. However it is obvious that, as in the classical case, this construction and all the results we present here, can be directly transposed mutatis mutandis to the case of an arbitrary underlying probability space with continuous filtration on it which satisfies the usual conditions.

Definition VI.4. *Let ν be a probability such that $\nu << \mu$. We note by $\mathcal{E}\left(\pi_., \mathcal{F}_.^\nu, \nu, H\right)$, the set of the simple process which is defined as the set of the $u \in L^0(\nu, H)$, which are of the form*

$$u(\omega) = \sum_{i=1}^{n-1} \alpha_i(\omega)\left(\pi_{t_{i+1}} - \pi_{t_i}\right)h_i$$

with $\alpha_i \in L^2\left(\mathcal{F}_{t_i}^\nu, \nu\right)$, $h_i \in H$ and $0 = t_1 < ... < t_n = 1$.

We will sometimes use the notation $\pi_{i;i+1} = \pi_{t_{i+1}} - \pi_{t_i}$.

Proposition VI.1. *With the notations of Definition VI.4 we have that $\mathcal{E}\left(\pi_., \mathcal{F}_.^\nu, \nu, H\right)$ is dense in $L_a^2(\nu, H)$*

Proof: Suppose that $v \in L_a^2(\nu, H)$ is orthogonal to $\mathcal{E}\left(\pi_., \mathcal{F}_.^\nu, \nu, H\right)$. Then for every $0 \leq s < t \leq 1$ and $a_s \in L_a^2\left(\mathcal{F}_s^\nu, \nu\right)$ and $l \in W^\star$ we have $E_\nu\left[< a_s\left(\pi_t - \pi_s\right)l, v >_H\right] = 0$. This means that $< \pi_. l, v >_H$ is a martingale on $(W, \mathcal{F}_.^\nu, \nu)$. On the other hand it is straightforward that $< \pi_. l, v >_H$ can be written as the difference between two increasing process, so that it vanishes as a martingale with finite variation. Hence $< l, v >_H = < \pi_1 l, v >_H = < \pi_0 l, v >_H = 0$ and $v = 0$. □

Let ν be a probability which is absolutely continuous with respect to μ. Further, assume that there is a $V \in L^0(\nu, W)$ such that $(V, H, \pi_.)$ is an abstract Wiener process on $(W, \mathcal{F}_.^\nu, \nu)$. For $u \in \mathcal{E}(\pi_., \mathcal{F}_.^\nu, \nu, H)$ the abstract stochastic integral of u with respect to V is defined by :

$$\delta^V u = \sum_{i=1}^{n-1} \alpha_i \delta^V(\pi_{i;i+1} h_i)$$

Obviously, the mapping $\delta^V : \mathcal{E}(\pi_., \mathcal{F}_.^\nu, \nu, H) \subset L_a^2(\nu, H) \to L^2(\nu)$ is an isometry. Since his domain is dense in $L_a^2(\nu, H)$, it extends as a uniformly continuous map, and this map is also an isometry $L_a^2(\nu, H) \to L^2(\nu)$ which we also note δ^V. For every $u \in L_a^2(\nu, H)$ $\delta^V u$ is called the abstract stochastic integral of u with respect to $(V, H, \pi_.)$. As in the classical case, this stochastic integral can be extended to a larger class, namely $L_{loc}^2(\nu, H)$.

Definition VI.5. *Let ν be a probability which is absolutely continuous with respect to μ. We define the set $L_{loc}^2(\nu, H)$ as the set of the adapted H- valued random variables u for which there exists an increasing sequence of $\mathcal{F}_.^\nu$ stopping times $(\tau_n)_{n\in\mathbb{N}}$, such that*

$$\nu\left(\lim_{n\to\infty} \tau_n = 1\right) = 1$$

and for every $n \in \mathbb{N}$

$$\pi_{\tau_n} u \in L_a^2(\nu, H)$$

For every u in $L_{loc}^2(\nu, H)$, and $(V, H, \pi_.)$ an $(W, \mathcal{F}_.^\nu, \nu)$ abstract Wiener process, it is easy to see that $\lim_{n\to\infty} \delta^V(\pi_{\tau_n} u)$ exists independently of the stopping time sequence we choose. Then one define univocally $\delta^V u = \lim_{n\to\infty} \delta^V(\pi_{\tau_n} u)$ for all such u. Moreover when $\nu = \mu$ and $V = I_W$ we will note $\delta^W u$ instead of $\delta^{I_W} u$ for $u \in L_{\text{loc}}^2(\mu, H)$

Proposition VI.2. *Let ν be a probability which is absolutely continuous with respect to μ and let $u \in L_{loc}^2(\mu, H)$. Further, assume that there is a $V \in L_a^0(\nu, W)$ of the form $V = I_W + v$, $v \in L_a^0(\nu, H)$ such that $(V, H, \pi_.)$ is an abstract Wiener process on $(V, \mathcal{F}_.^\nu, \nu)$. Then we have*

$$\delta^V u = \delta^W u + < v, u >_H$$

$\nu - a.s.$

Proof: The claim is obvious when u is a step process since $\nu << \mu$ and a fortiori $\mathcal{F}_.^\mu \subset \mathcal{F}_.^\nu$. The generalization is straightforward. □

Suppose now that for a $v, b \in L_a^0(\nu, H)$, $(V = I_W + v, H, \pi_.)$ and $(B = I_W + b, H, \pi_.)$ are two abstract Wiener process on $(W, \mathcal{F}_.^\nu, \nu)$. Then, it is easy to see that for any simple process $u \in \mathcal{E}(\pi_., \mathcal{F}_.^\nu, \nu, H)$ we have $\nu - a.s.$, $\delta^V u - < v, u >_H = \delta^B u - < b, u >_H$ and to extend this property to $L_a^0(\nu, H)$. Hence the following definition, which is suggested by the last proposition, is unequivocal

Definition VI.6. *Let ν be a probability which is absolutely continuous with respect to μ and let $u \in L_a^0(\nu, H)$. Further, assume that there is a $V \in L_a^0(\nu, W)$ of the form $V = I_W + v$, $v \in L_a^0(\nu, H)$ such that $(V, H, \pi_.)$ is an abstract Wiener process on $(W, \mathcal{F}_.^\nu, \nu)$. Then we note $\delta^W u$ the element of $L^0(\nu)$ defined by the relation*

$$\delta^W u = \delta^V u - < v, u >_H$$

$\nu - a.s.$

The following property is the key of almost all our further calculations.

Proposition VI.3. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ, and let $u \in L^2_{loc}(\nu, H)$. Further, assume that there is a $V \in L^0(\nu, W)$ such that $(V, H, \pi_.)$ is an abstract Wiener process on $(W, \mathcal{F}^\nu_., \nu)$. Then the process $(t, w) \to \delta^V(\pi_t u) = m_t(w)$ is a continuous local martingale on $(W, \mathcal{F}^\nu_., \nu)$. Moreover, the associated increasing process is such that for every $t \in [0,1]$*

$$< \delta^V(\pi_. u) >_t = |\pi_t u|^2_H$$

In particular, when u is an element of $L^2_a(\mu, H)$, $\delta^W(\pi_. u)$ is a martingale.

Proof: The result is straightforward in the case where u is a simple process. By a standard limiting procedure, the results holds for u in $L^2_a(\nu, H)$, and by usual stopping techniques it also holds when for any $u \in L^2_{loc}(\nu, H)$. When u is in $L^2_a(\nu, H)$, since for every $t \in [0,1]$

$$E_\nu\left[\left\langle \delta^V(\pi_. u)\right\rangle_t\right] = E_\nu\left[|\pi_t u|^2_H\right] \leq E_\nu\left[|u|^2_H\right]$$

we get by a classical result of stochastic calculus that $\delta^V(\pi_. u)$ is a martingale. □

The proof of the next proposition, uses an integration over a vector valued measures. Since we won't use this notion in the sequel, we won't neither define it here and refer to **[50]** p42-44. for much details about the notations we are using in the next proof.

Proposition VI.4. *Let ν be a probability which is absolutely continuous with respect to μ. Moreover, assume that $(V, H, \pi_.)$ is an abstract Wiener process on $(W, \mathcal{F}^\nu_., \nu)$. Let $u \in L^2_{loc}(\nu, H)$ and $(H_s, s \in [0,1])$ be a real valued measurable $\mathcal{F}^\nu_.$ adapted random function such that*

$$\int_0^1 H_s d < \pi_s u, u >_H < \infty$$

$\nu-$a.s. Then there is a $\eta \in L^2_{loc}(\nu, H)$ such that

$$\int_0^{.} H_s d\delta^V \pi_s u = \delta^V(\pi_. \eta)$$

in particular, for all $\alpha \in L^2_{\mathrm{loc}}(\nu, H)$, $\left\langle \delta^V \pi_. \alpha, \delta^V \pi_. \eta\right\rangle_t = \int_0^t H_s d < \alpha, \pi_s u >_H$

Proof: The first point can be obtained by approximating H with simple adapted functions, for which the result is obvious, and then extended to $L^2_a(\nu, H)$ and then to $L^2_{loc}(\nu, H)$. We recall that if M and N are two local martingale and H is integrable with respect to M, by a characteristic property of the stochastic integral we have $\left\langle \int_0^{.} H_s dM_s, N_. \right\rangle_t = \int_0^t H_s d < M, N >_t$. Hence by taking $M_. = \delta^V \pi_. u$ and $N_. = \delta^V \pi_. \alpha$:

$$\begin{aligned}
\left\langle \delta^V \pi_. \eta, \delta^V \pi_. \alpha \right\rangle_t &= \left\langle \int_0^{.} H_s d\delta^V \pi_s u, \delta^V \pi_. \alpha \right\rangle_t \\
&= \int_0^t H_s d \left\langle \delta^V \pi_. u, \delta^V \pi_. \alpha \right\rangle_s \\
&= \int_0^t H_s d < \pi_s u, \alpha >_H
\end{aligned}$$

□

With this proposition in hand, the proof of the following theorem is almost the same as in the classical case (see **[50]** for a detailed proof).

Theorem VI.1. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ and $(V, \pi_., H)$ be an abstract Wiener process on (W, ν) with respect to its own filtration $\mathcal{F}^V_.$. Moreover, assume that M is a continuous abstract local martingale on $\left(W, \mathcal{F}^V_., \nu\right)$ and $M_0 = 0$. Then there is a $\alpha \in L^2_{loc}(\nu, H)$ such that*

$$M_. = -\delta^V \pi_. \alpha$$

2.3. The Monge-Kantorovich problem on Wiener space. The Monge and Monge-Kantorovich problems on abstract Wiener space have been treated for instance in [**12**] or [**13**]. Some of these results are summed up and used to get results in the classical Wiener space in [**47**]. In order to get the proof of the results we present here and a more complete overview on this topic, we refer to these three articles and to the references therein, in particular to [**55**]. In this section we only recall the main results that we will use in further sections. We recall that the Wasserstein distance was defined in Definition I.1. In the sequel we shall relate the invertibility of adapted perturbations of the identity to the existence of some variational formulations of the Monge-Kantorovitch problem in the Gaussian case. We recall the following result which is a particular case of the Theorem 3.2 of [**13**].

Theorem VI.2. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ. Assume that $d(\nu,\mu) < \infty$ (see Definition I.1), then there is a measurable mapping $T : W \to W$ which is solution to the original Monge problem. Moreover its graph supports the unique solution of the Monge-Kantorovitch problem γ i.e.*

$$(I_W \times T)\,\mu = \gamma$$

in particular $T\mu = \nu$, $T - I_W \in L^2(\mu, H)$, and there is a mapping T^{-1} such that

$$\mu\left(\left\{\omega | T^{-1} \circ T = I_W\right\}\right) = \nu\left(\left\{\omega | T \circ T^{-1} = I_W\right\}\right) = 1$$

3. The Girsanov shift on abstract Wiener space

In this section we provide a sharp version of the Girsanov theorem on abstract Wiener space and we extend some results of [**14**] and [**15**] to our framework. The Lemma VI.1 will be used to prove Theorem VI.3. This latter theorem is very useful and we shall apply it implicitly whenever we will meet a probability ν which is absolutely continuous with respect to the Wiener measure μ. Part of this result already appears in [**50**] (Proposition 2.5.1) but with a wrong proof. As in [**50**] the main idea of the proof of Theorem VI.3 is still to apply the theorem of representation. Here we correct the proof and avoid a vicious circle, by first showing the existence of an abstract Wiener process on $(W, \mathcal{F}^{\nu}_{.}, \nu)$ thanks to Lemma VI.1.

Lemma VI.1. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ and let N be a semimartingale on $\left(W, \mathcal{F}^{\nu}, (\mathcal{F}^{\nu}_t)_{t\in[0,1]}, \nu\right)$. Then there is a $v \in M_{\nu}\left((W, \mathcal{F}^{\nu}), (W, \mathcal{B}(W))\right)$ such that*

- *$\nu -as$ for every $t \in [0,1]$ and for every $l \in W^{\star}$, $< N, \delta^W \pi_{.} l >_t = < \pi_t l, v >_{W^{\star},W}$*
- *$v \in L^0_a(\nu, H)$ i.e. v is adapted to the $\nu-$augmented canonical filtration $\mathcal{F}^{\nu}_{.} = \sigma\left(\mathcal{F}^{W,0}_{.} \cup \Theta^{\nu}\right)$ and $\nu - a.s.$ $v \in H$*

Proof: By the classical Girsanov theorem we know that for any $h \in H$, the $\mu-$martingale (i.e. martingale on $(W, \mathcal{F}^{\mu}_{.}, \mu)$) $\delta^W \pi_{.} h$ is also a $\nu-$semimartingale (i.e. a semimartingale on $(W, \mathcal{F}^{\nu}_{.}, \nu)$), so that the process $(t,\omega) \to < \delta^W \pi_{.} h, N >_t (\omega)$ is well defined. We set

$$\begin{array}{cccc} u & : & H & \to & L^0(\nu) \\ & & h & \mapsto & < \delta^W \pi_{.} h, N >_1 \end{array}$$

Let $(h_i)_{i\in\mathbb{N}}$ be an orthonormal basis of H, we note π_n the projection on the vector space spanned by $(h_1, ..., h_n)$, and $H_n = \pi_n(H)$. We note j_n the canonical bijection such that $j_n(\mathbb{R}^n) = H_n$, and we note i_n the inverse of j_n. We define

$$\begin{array}{cccc} \widehat{u_n} & : & \mathbb{R}^n & \to & L^0(\nu) \\ & & x & \mapsto & < \delta^W \pi_{.} j_n(x), N >_1 \end{array}$$

so that we have

$$|\widehat{u_n}(x) - \widehat{u_n}(y)| \leq \sqrt{< N, N >_1} |x - y|_{\mathbb{R}^n}$$

$\nu - a.s.$ where $< N, N >_1$ is finite $\nu - as$, and where the set on which the last inequality fails may depend on x and y. Hence for every $\gamma > 0$, there is a probability ν_γ equivalent to ν such that

$$E_{\nu_\gamma}\left[|\widehat{u_n}(x) - \widehat{u_n}(y)|^\gamma\right] \leq C_\gamma |x-y|^\gamma_{\mathbb{R}^n}$$

where C_γ only depends on N and γ. Thus we can apply the Kolmogorov lemma to get a modification $\overline{u_n}$ of $\widehat{u_n}$ with the property that $\nu-$ a.s., the map $x \to \overline{u_n}(\omega, x)$ is continuous. We set $u_n = \overline{u_n}$ on the set where $x \to \overline{u_n}(\omega, x)$ is continuous, and $u_n = 0$ otherwise. Therefore $x \to u_n(\omega, x)$ is continuous for every $\omega \in W$, and $u_n(x) =< \delta^W \pi . j_n(x), N >_1 \nu - a.s.$ where the set on which the equality fails may depend on $x \in \mathbb{R}^n$. As a matter of fact, the continuity will enable us to choose a linear modification of u_n. For every n-uplet of rationals $(\lambda_1, ..., \lambda_n) \in \mathbb{Q}^n$, let

$$\Omega_{\lambda_1,...,\lambda_n} = \left\{\omega \middle| u_n\left(\sum_{i=1}^n \lambda_i e_i\right) = \sum_{i=1}^n \lambda_i u_n(e_i)\right\}$$

where $(e_1, ..., e_n)$ is an orthonormal basis of $\mathbb{R}^n$. By continuity of u_n,

$$\bigcap_{\lambda_1 \in \mathbb{Q}, ..., \lambda_n \in \mathbb{Q}} \Omega_{\lambda_1,...,\lambda_n} = \Omega_l$$

where

$$\Omega_l = \left\{\omega \middle| \forall x \in \mathbb{R}^n \ u_n(x) = \sum_{i=1}^n < x, e_i >_{\mathbb{R}^n} e_i\right\}$$

Moreover by definition of $< .,. >_1$, $\nu(\Omega_{\lambda_1,...,\lambda_n}) = 1$ so that $\nu(\Omega_l) = 1$. We also set

$$\Omega_f = \{\omega | u_n(e_i) < \infty \ \forall i \in [1...n]\}$$

By construction of $< .,. >_1$, we still have $\nu(\Omega_f) = 1$. We set for every $h \in H$ and every $\omega \in W$

$$f_n(\omega, h) = \begin{cases} u_n(\omega, i_n \circ \pi_n h) & if \qquad \omega \in \Omega_l \cap \Omega_f, \\ 0 & otherwise \end{cases}$$

By the continuity and the linearity of i_n and π_n, $h \in H \to f_n(\omega, h) \in \mathbb{R}$ inherit both the linearity and the continuity of u_n on every $\omega \in \Omega_l \cap \Omega_f$, so that it is both linear and continuous on the whole space W. Hence we can define $\widehat{f_n} : \omega \in W \to H^\star$ where $\widehat{f_n}(\omega) : h \in H \to f_n(\omega, h)$. However we can't be sure to have $|\widehat{f_n}|_{H^\star} \leq \sqrt{< N, N >_1} \ \nu - a.s$, and we will have to modify it. For that reason, let $(k_i)_{i \in \mathbb{N}}$ be a dense subset of H. For every k_i, we set:

$$\widetilde{\Omega}_{k_i} = \left\{\omega \middle| |f_n(\omega, k_i)| \leq \sqrt{< N, N >_1}(\omega) |k_i|_H\right\}$$

By definition of f_n, $\nu\left(\widetilde{\Omega}_{k_i}\right) = 1$. Moreover, by the continuity of $h \in H \to f_n(\omega, h) \in \mathbb{R}$, we also have

$$\bigcap_{i \in \mathbb{N}} \widetilde{\Omega}_{k_i} = \Omega_m$$

where

$$\Omega_m := \{\omega | f_n(\omega, h)| \leq < N, N >_1 (\omega) |h|_H, \forall h \in H\}$$

Thus $\nu(\Omega_m) = 1$. We can now define a modification $\widetilde{f_n}$ of f_n by :

$$\widetilde{f_n}(\omega, h) = \begin{cases} f_n(\omega, h) & if \qquad \omega \in \Omega_m, \\ 0 & otherwise \end{cases}$$

By construction, for every $\omega \in W$, $h \to \widetilde{f_n}(\omega, h)$ is in $H^\star$ with a norm in $H^\star$ which is less than $\sqrt{< N, N >}(\omega)$. Finally for every $\omega \in W$, we note $v_n(\omega)$ the element of H which is associated, through the Riesz representation theorem, with the element of $H^\star$, $h \to \widetilde{f_n}(\omega, h)$. Still by construction, we have for every $\omega \in W$ for every $h \in H$ $< h, v_n(\omega) >_H = u_n(\omega, i_n \circ \pi_n h)$, if $\omega \in \Omega_l \cap \Omega_f \cap \Omega_m$, and $< h, v_n >_H = 0$ otherwise. Hence we have $\pi_m v_n = v_m$ for every $m \leq n$. Together with the fact that $|v_n|_H \leq \sqrt{<N,N>_1}$, and since there is a $K>0$ such that $|.|_W \leq K|.|_H$, we have

$$\begin{aligned}\nu\left(|v_n - v_m|_W > c\right) &\leq \nu\left(K|v_n - v_m|_H > c\right)\\ &= \nu\left(K|\left(\pi_n - \pi_m\right)v_n|_H > c\right)\\ &\leq \nu\left(K\sqrt{<N,N>_1}|\pi_n - \pi_m|_{op} > c\right)\\ &\to 0 \; n,m\to\infty\end{aligned}$$

We note $v \in M_\nu\left(\left(W,\mathcal{F}^\nu\right),\left(W,\mathcal{B}\left(W\right)\right)\right)$ the limit in probability of $\left(v_n\right)_{n\in\mathbb{N}}$. By construction we have $|v|_H \leq \sqrt{<N,N>_1}$, so that $\nu - a.s.$ $v \in H$. On the other hand, still by construction, it is easy to see that $u_n\left(\omega, i_n \circ \pi_n\right)$ converges to u under the probability ν. Therefore $\nu - a.s$ for every $h \in H$ $<v,h>_H = u\left(h\right) = <\delta^W h, N>_1$. Furthermore, we remark that, $<N, \delta^W \pi_. h>_t = <N, \delta^W \pi_{.\wedge t} h>_1 = <N, \delta^W \pi_. \pi_t h>_1$ $\nu - a.s.$, so that for every $t \in [0,1]$, ν almost surely we have $<N, \delta^W \pi_. h>_t = <\pi_t h, v>_H$. in particular, let, $l \in W^\star$, $B \in \mathcal{B}\left(\mathbb{R}\right)$, and

$$\Omega = \left\{\omega \middle| <\pi_t l, v>_H = <N, \delta^W \pi_. l>_t\right\}$$

Then,

$$\left(<\pi_t l, v>_H\right)^{-1}\left(B\right) = \left(\Omega \cap \left[\left(<\pi_t l, v>_H\right)^{-1}\left(B\right)\right]\right) \cup \left(\Omega^c \cap \left[\left(<\pi_t l, v>_H\right)^{-1}\left(B\right)\right]\right)$$

On the one hand, $\left(<\pi_t l, v>_H\right)^{-1}\left(B\right) \cap \Omega^c \subset \Omega^c \in \Theta^\nu$. On the other hand $\left(<\pi_t l, v>_H\right)^{-1}\left(B\right) \cap \Omega \subset \left(<N, \delta^W \pi_. l>_t\right)^{-1}\left(B\right) \subset \mathcal{F}_t^\nu$. Hence $\left(<\pi_t l, v>_H\right)^{-1}\left(B\right) \subset \mathcal{F}_t^\nu$. This shows that $<\pi_t l, v>_H$ is $\mathcal{F}_t^\nu$ measurable for every t. Note that the fact that v is adapted doesn't depend on the modification we choose, since $\mathcal{F}_t^\nu$ is complete with respect to ν. □

Theorem VI.3. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ ($\nu << \mu$). Then there is a unique $v \in L_a^0\left(\nu, H\right)$ such that $\left(V = I_W + v, H, \pi_.\right)$ is an abstract Wiener process on $\left(W, \mathcal{F}^\nu, \left(\mathcal{F}_t^\nu\right)_{t\in[0,1]}, \nu\right)$ with the property that $\nu - a.s.$*

$$E_\mu\left[\frac{d\nu}{d\mu}\middle|\mathcal{F}_t^\mu\right] = \exp\left(-\delta^W \pi_t v - \frac{|\pi_t v|_H^2}{2}\right)$$

for every $t \in [0,1]$, where the notation $\delta^W v$ is defined in Definition VI.6. We call $V = I_W + v$ (resp. v) the Girsanov shift (resp. drift) associated with ν.

Proof: In this proof we note $L = \frac{d\nu}{d\mu}$ the associated Radon-Nykodym density and $L_t = E_\mu\left[L|\mathcal{F}_t^\mu\right] = \frac{d\nu|_{\mathcal{F}_t^\mu}}{d\mu|_{\mathcal{F}_t^\mu}}$. Since ν is a probability measure we have $L_0 = 1$, and $L_. = E_\mu\left[L|\mathcal{F}_.^\mu\right]$ is a martingale on $\left(W, \mathcal{F}_.^\mu, \mu\right)$. Let $\tau_0 = inf\left\{t : L_t = 0\right\}$ if $\left\{t : L_t = 0\right\} \neq \emptyset$, and $\tau_0 = \infty$ otherwise. By applying Doob's optional stopping theorem we have :

$$\begin{aligned}\nu\left(\left\{\omega|\tau_0 < \infty\right\}\right) &= E_\mu\left[L_1 1_{\tau_0\in[0,1]}\right]\\ &= E_\mu\left[L_{\tau_0} 1_{\tau_0\in[0,1]}\right]\\ &= 0\end{aligned}$$

Hence, the random process $\left(\omega, t\right) \to \ln L_t\left(\omega\right)$ is well defined with values in $\mathbb{R}$ ν a.s. From the classical Girsanov theorem we know that to every local martingale M on $\left(W, \mathcal{F}_.^\mu, \mu\right)$ we can associate a Girsanov transform $\widetilde{M}$ defined by $\widetilde{M} = M + L_.^{-1}. <L_., M>$ (see for instance [**7**]) , which is such that $\widetilde{M}_.$ is a local martingale on $\left(W, \mathcal{F}_.^\nu, \nu\right)$. In particular every semimartingales on $\left(W, \mathcal{F}_.^\mu, \mu\right)$, is also semimartingales on $\left(W, \mathcal{F}_.^\nu, \nu\right)$. Therefore the Ito formula applied to $\left(t, \omega\right) \to \ln\left(L_t\right)\left(t, \omega\right)$ on $\left(W, \mathcal{F}_.^\nu, \nu\right)$ yields

$$dln\left(L_t\right) = \frac{dL_t}{L_t} - \frac{d<L., L.>_t}{2L_t^2}$$

Thus, if we set $N_. = -\int_0^{\cdot} \frac{dL_t}{L_t}$, $N_.$ is a semimartingale on $(W, \mathcal{F}^{\nu}_{.}, \nu)$ such that $\nu - a.s.$ for every $t \in [0,1]$

$$L_t = \exp\left(-N_t - \frac{1}{2} < N, N >_t\right)$$

Let $l \in W^{\star}$, since $\delta^W \pi_. l$ is a local martingale on $(W, \mathcal{F}^{\mu}_{.}, \mu)$, we know that its Girsanov transform

$$\widetilde{\delta\pi_. l_.} = \delta^W \pi_. l + < N, \delta^W \pi_. l >.$$

is a $(W, \mathcal{F}^{\nu}_{.}, \nu)$ local martingale. Hence, if we note $v \in L^0_a(\nu, H)$ the H-valued and adapted random variable associated with $N_.$ in Lemma VI.1, and $V = I_W + v$, we have for every $t \in [0,1]$

$$\begin{aligned}
\widetilde{\delta\pi_. l_t} &= \delta^W \pi_t l + < N, \delta^W \pi_. l >_t \\
&= \delta^W \pi_t l + < \pi_t l, v >_{W^{\star}, W} \\
&= < \pi_t l, V >_{W^{\star}, W}
\end{aligned}$$

which shows that V is an abstract martingale on $(W, \mathcal{F}^{\nu}_{.}, \nu)$. Moreover, $< \widetilde{\delta\pi_. l} >_t = < \delta^W \pi_. l >_t = |\pi_t l|^2_H$ so that $(V, H, \pi_.)$ is an abstract Wiener process on $(W, \mathcal{F}^{\nu}_{.}, \nu)$. The existence of this Wiener process will enable us to use the result of the previous section. Since L_t is a martingale on $(W, \mathcal{F}^{\mu}_{.}, \mu)$, by Theorem VI.1 there is a $\alpha \in L^2_{loc}(\mu, H)$ such that $L_. = -\delta^W \pi_. \alpha$, and thus $N_t = \int_O^t \frac{d\delta^W \pi_. \alpha}{L_t}$. By using Proposition VI.2, it implies that

$$N_t = \int_0^t \frac{d\delta^V \pi_s \alpha}{L_s} - \int_0^t \frac{d < v, \pi_s \alpha >_H}{L_s}$$

On the other hand, the Proposition VI.4 (with $H_s = \frac{1}{L_s}$) shows the existence of a $\eta \in L^2_{loc}(\nu, H)$ such that $N_t = \delta^V \pi_t \eta - < \delta^V \pi_. \eta, \delta^V \pi_. v >_t = \delta^V \pi_t \eta - < \pi_t \eta, v >_H$ $\nu - a.s.$ Hence for $l \in W^{\star}$, we have $\nu - a.s.$

$$< \pi_t v, l >_H = \left\langle N_., \delta^W \pi_. l \right\rangle_t = \left\langle \delta^V \pi_. \eta, \delta^V \pi_. l \right\rangle_t = < \pi_t \eta, l >_H$$

where the set on which the identity fails may depend on l. Let $(l_i)_{i \in \mathbb{N}}$ be a dense subset of $W^{\star}$. For every l_i, we set : $\widetilde{\Omega}_{l_i} = \{\omega | < \eta, l_i >_H = < v, l_i >_H\}$ so that $\nu\left(\widetilde{\Omega}_{l_i}\right) = 1$. By the continuity of the scalar product we also have

$$\cap_{i \in \mathbb{N}} \widetilde{\Omega}_{l_i} = \Omega_e$$

where

$$\Omega_e := \{\omega | < \eta, l >_H = < v, l >_H, \forall l \in W^{\star}\}$$

Hence $\nu(\Omega_e) = 1$ which means that $v = \eta$ $\nu - a.s$ so that

$$N_t = \delta^V \pi_t v - < \pi_t v, v > = \delta^W \pi_t v$$

$\nu - a.s.$ Thus, $\nu - a.s.$ for every $t \in [0,1]$

$$L_t = \exp\left(-\delta^W \pi_t v - \frac{|\pi_t v|^2_H}{2}\right)$$

and by taking $t = 1$ we get the result (the proof of the uniqueness can be achieved the same way). Finally, we prove the measurablity :

$$\begin{aligned}
V^{-1}(\mathcal{F}^{\mu}) &= V^{-1}\left(\sigma\left(\mathcal{F}^{W,0} \cup \Theta^{\mu}\right)\right) \\
&= \sigma\left(V^{-1}\left(\mathcal{F}^{W,0} \cup \Theta^{\mu}\right)\right) \\
&= \sigma\left(V^{-1}\left(\mathcal{F}^{W,0}\right) \cup V^{-1}(\Theta^{\mu})\right) \\
&\subset \sigma\left(\mathcal{F}^{\nu} \cup V^{-1}(\Theta^{\mu})\right)
\end{aligned}$$

Where the last line comes from the fact that $v \in \mathcal{M}((W, \mathcal{F}^{\nu}), (W, \mathcal{B}(W)))$ and $\mathcal{B}(W) = \mathcal{F}^{W,0}$. On the other hand let $N \in \Theta^{\mu}$, then there exist a $A \in \mathcal{B}(W)$ such that $N \subset A$ and $\mu(A) = 0$. We have

$0 = \mu(A) = V\nu(A) = \nu\left(V^{-1}(A)\right)$. Hence $V^{-1}(N)$ which is contained in the $\nu-$ null set $V^{-1}(A)$, is in Θ^ν so that $V^{-1}(\Theta^\mu) \subset \Theta^\nu$. Finally $V^{-1}(\mathcal{F}^\mu) \subset \mathcal{F}^\nu$. This prove that $V \in \mathcal{M}(\mathcal{F}^\nu, \mathcal{F}^\mu)$. □

Henceforth for any probability ν which is absolutely continuous with respect to μ and for any $v \in L_a^0(\nu, H)$ we set the following notation which will be used throughout the paper :

$$\rho\left(-\delta^W v\right) := \exp\left(-\delta^W v - \frac{|v|_H^2}{2}\right)$$

Corollary VI.1. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ. Then we have*

$$\mu \sim \nu$$

if and only if there is a $v \in L_a^0(\mu, H)$ such that

$$\frac{d\nu}{d\mu} = \rho\left(-\delta^W v\right)$$

$\mu - as$ and in that case $V = I_W + v$ is the Girsanov shift associated with ν

Proof: Since $\mu \sim \nu$ if and only if $\frac{d\nu}{d\mu} > 0$ $\mu - as$, this is a straightforward consequence of Theorem VI.3. □

Proposition VI.5 is well known in the classical case (see lemma 2.6 of **[15]** and Proposition 2.11 of **[14]**) : we omit the proof which is very similar.

Proposition VI.5. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ, and let $V = I_W + v$ denote the Girsanov shift of ν. Then we have*

$$2H(\nu|\mu) = E_\nu\left[|v|_H^2\right]$$

In particular

$$v \in L^2(\nu, H)$$

if and only if

$$H(\nu|\mu) < \infty$$

As a first consequence we get the following Talagrand inequality (see **[42]** and **[45]** where this nice proof seems to have appeared for the first time in the case of probabilities equivalent to the Wiener measure on the classical Wiener space).

Proposition VI.6. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ. Then*

$$d(\nu, \mu) \leq 2H(\nu|\mu)$$

Proof: Because of Proposition VI.5, the Girsanov shift V associated with ν is such that

$$2H(\nu|\mu) = E_\nu\left[|V - I_W|_H^2\right]$$

On the other hand by Theorem VI.3 we have $(V \times I_W)\nu \in \Sigma(\mu, \nu)$. Therefore $d(\nu, \mu) \leq E_\nu\left[|V - I_W|_H^2\right]$ and we get the result. □

4. Pull-back of the Girsanov shift with related morphisms

4.1. Some basic facts.

Definition VI.7. *Let P and Q be two probabilities on $(W, \mathcal{B}(W))$ which are absolutely continuous with respect to the Wiener measure. Denote by $\mathcal{F}^P$ (resp. by $\mathcal{F}^Q$) the completed sigma field of $\mathcal{B}(W)$ with respect to P (resp. Q), and still note by P (resp. Q) the unique extension of P on $\mathcal{F}^P$ (resp. of Q on $\mathcal{F}^Q$). Then we note $\mathcal{R}(P,Q)$ the set of the morphisms of probability spaces between $\left(W, \mathcal{F}^P, P\right)$ and $\left(W, \mathcal{F}^Q, Q\right)$ which is defined by :*

$$\mathcal{R}(P,Q) = \left\{U \in M_P\left(\left(W, \mathcal{F}^P\right), \left(W, \mathcal{F}^Q\right)\right) : UP = Q\right\}$$

in particular :

- *$\mathcal{R}(P,P)$ is called the set of the $P-$ rotations*
- *When $U \in \mathcal{R}(\mu,\nu)$ where μ is the Wiener measure and $\nu << \mu$, we will say that U represents ν. Moreover we will note $\mathcal{R}_a(\mu,\nu)$ the subset of the $U \in \mathcal{R}(\mu,\nu)$ such that U is an API, i.e. of the form $U = I_W + u$ with $u \in L_a^0(\mu, H)$*
- *Similarly we note $\mathcal{R}_a(\nu,\mu)$ the subset of the $V \in \mathcal{R}(\nu,\mu)$ such that V is an $\nu - API$, i.e. of the form $V = I_W + v$ with $v \in L_a^0(\nu, H)$*

It is routine to check the following facts. Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ, $U \in \mathcal{L}^0(\mu, W)$ and $V \in \mathcal{L}^0(\nu, W)$. We further assume that $U\mu << \nu$ (resp. $V\nu << \mu$). Then $V \circ U \in \mathcal{L}^0(\mu, W)$ (resp. $U \circ V \in \mathcal{L}^0(\nu, W)$) and the equivalence class of $V \circ U$ in $L^0(\mu, W)$ (resp. of $U \circ V$ in $L^0(\nu, W)$) only depends (in both cases) on the equivalence class of U in $L^0(\mu, W)$ and of V in $L^0(\nu, W)$. As a consequence, if $U \in L^0(\mu, W)$ and $V \in L^0(\nu, W)$ are such that $U\mu << \nu$ and $V\nu << \mu$, then $V \circ U$ and $U \circ V$ are well defined as above and we have $V \circ U\mu << \mu$ and $U \circ V\nu << \nu$ (in particular this implies that $U \circ V \in M_\nu((W, \mathcal{F}^\nu), (W, \mathcal{F}^\nu))$). Similarly when $v = V - I_W \in L^0(\nu, H)$ (resp. $v = V - I_W \in L_a^0(\nu, H)$), and $U \in L^0(\mu, W)$ such that $U\mu << \nu$ it is easy to see that $v \circ U \in L^0(\mu, H)$ (resp. that $v \circ U \in L^0(\mu, H)$ is adapted with respect to the augmented filtration generated by U). The following fact which is trivial has to be noticed :

Proposition VI.7. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure. Then*

$$\mathcal{R}(\mu,\nu) = \left\{U \in L^0(\mu, W) : U\mu = \nu\right\}$$

Moreover for every $(U,V) \in \mathcal{R}(\mu,\nu) \times \mathcal{R}(\nu,\mu)$, we have $V \circ U \in \mathcal{R}(\mu,\mu)$ and $U \circ V \in \mathcal{R}(\nu,\nu)$

4.2. Some Inverse images associated with the Girsanov shift. We prove two theorems, and a proposition. Theorem VI.4, already appeared for instance in **[47]** in the case of the adapted perturbations of the identity : it deals with the pullback of morphisms of probability space under a related Girsanov shift. On the contrary, Theorem VI.5 deals with pullbacks of the Girsanov shift under associated morphisms of probability space. Within our framework, the proof of Theorem VI.4 is trivial, and the proof of Theorem VI.5 is very straightforward. However Theorem VI.5 has many important consequences.

Theorem VI.4. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ, and let $V = I_W + v \in \mathcal{R}_a(\nu,\mu)$ be the Girsanov shift associated with ν. For every $X \in L^0(\mu, W)$, we define $L^X \in L^0(\nu, W)$ by*

$$L^X = X \circ V$$

Then X represents ν (i.e. $X\mu = \nu$) if and only if $L^X \in \mathcal{R}(\nu,\nu)$. Moreover in that case $\left(L^X, H, \pi_.\right)$ is an abstract martingale on $\left(W, \mathcal{F}_.^{L^X}, \nu\right)$ if and only if $(X, H, \pi_.)$ is an abstract martingale on $\left(W, \mathcal{F}_.^X, \mu\right)$

Proof: The Theorem VI.3 implies in particular that $V\nu = \mu$ so that we also have $L^X\nu = X\mu$. Therefore $L^X\nu = \nu$ if and only if $X\mu = \nu$. We focus on the second part of the proof which shows under which condition the triplet $\left(L^X, H, \pi_.\right)$ is an abstract martingale on $\left(W, \mathcal{F}_.^{S^X}, \nu\right)$. Let l be in $W^\star$. Then for all $s < t$ and for all θ which is $\mathcal{F}_s^\nu$ measurable, the Theorem VI.3 yields

$$E_\nu\left[< (\pi_t - \pi_s)\, l, L^X > \theta \circ L^X\right] \quad = \quad E_\mu\left[< (\pi_t - \pi_s)\, l, X > \theta \circ X\right]$$

From which the result follows directly. □

The next theorem is well known in the case of the classical Wiener space (at least in the case of measures equivalent to the Wiener measure), and is very useful to study the quasi-invariant flows on Wiener space. Within this framework, our proof is straightforward, and the result is very general : for instance it has to be compared with the Proposition 2.5.1 of **[50]** in the classical case. As we shall see later, as long as we still consider stochastic processes as morphisms of probability spaces, many known and unknown applications can be obtained more or less directly from this result.

Theorem VI.5. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ. Let $V = I_W + v \in \mathcal{R}_a(\nu, \mu)$ denote the Girsanov shift associated with ν and let $X \in \mathcal{R}(\mu, \nu)$ (i.e. X represents ν). If we note $R^X = V \circ X$, then $\left(R^X, H, \pi_.\right)$ is a $\left(W, \mathcal{F}_.^X, \mu\right)$ abstract Wiener process, where $\mathcal{F}_.^X = \sigma\left(\mathcal{F}_.^{X,0} \cup \Theta^\mu\right)$ is the augmented filtration of X.*

Proof: Let $R^X = V \circ X$ which is a rotation by Theorem VI.3 ($R^X\mu = V \circ X\mu = V\nu = \nu$), and let l be in $W^\star$. Since V is adapted to $\mathcal{F}_.^\nu$, $\delta^{R^X}\pi_. l =< \pi_. l, R^X >_{W^\star, W}$ is adapted to the augmented filtration associated with X, $\mathcal{F}_.^X = \sigma\left(\mathcal{F}_.^{X,0} \cup \Theta^\mu\right)$. Indeed for every $l \in W^\star$ and $B \in \mathcal{B}(\mathbb{R})$ and $t \in [0,1]$, we have

$$\begin{aligned}
\left(< \pi_t l, R^X >\right)^{-1}(B) &= \left(< \pi_t l, V \circ X >\right)^{-1}(B) \\
&= X^{-1}\left(\left(< \pi_t l, V >\right)^{-1}(B)\right) \\
&\subset X^{-1}\left(\mathcal{F}_t^\nu\right) \\
&= X^{-1}\left(\sigma\left(\mathcal{F}_t^{W,0} \cup \Theta^\nu\right)\right) \\
&= \sigma\left(X^{-1}\left(\mathcal{F}_t^{W,0}\right) \cup X^{-1}\left(\Theta^\nu\right)\right) \\
&= \sigma\left(\mathcal{F}_t^{X,0} \cup X^{-1}\left(\Theta^\nu\right)\right)
\end{aligned}$$

Let $N \in \Theta^\nu$, then there is a $A \in \mathcal{B}(W)$ such that $\nu(A) = 0$ and $N \subset A$. Since $X\mu = \nu$ we have $\mu\left(X^{-1}(A)\right) = \nu(A) = 0$. Hence $X^{-1}(N) \subset X^{-1}(A)$ is in Θ^μ, so that $X^{-1}(\Theta^\nu) \subset \Theta^\mu$ Thus $\mathcal{F}_.^{R^X} \subset \mathcal{F}_.^X = \sigma\left(\mathcal{F}_.^{X,0} \cup \Theta^\mu\right)$ i.e. R^X is adapted to the augmented filtration generated by X. Moreover, let s and t be such that $s < t$, and let θ be a bounded and $\mathcal{F}_s^\nu$ measurable mapping, we have

$$\begin{aligned}
E_\mu\left[\left(\delta^{R^X}\pi_t l - \delta^{R^X}\pi_s l\right)\theta \circ X\right] &= E_\mu\left[< \pi_t l - \pi_s l, R^X > \theta \circ X\right] \\
&= E_\mu\left[< \pi_t l - \pi_s l, V \circ X > \theta \circ X\right] \\
&= E_\nu\left[\left(\delta^V \pi_t l - \delta^V \pi_s l\right)\theta\right] \\
&= 0
\end{aligned}$$

Where the last equality is a straightforward consequence of Theorem VI.3 which states that $(V, H, \pi_.)$ is an abstract Wiener process on $(W, \mathcal{F}_.^\nu, \nu)$. Thus $\delta^R\pi_. l$ is a $\left(X, \mathcal{F}_.^X, \mu\right)$ martingale. Moreover, with the same notations we have :

$$\begin{aligned}
E_\mu\left[\left(\left(\delta^{R^X}\pi_t l\right)^2-|\pi_t l|_H^2\right)\theta\circ X\right] &= E_\mu\left[\left(\left(<\pi_t l, R^X>_{W^\star,W}\right)^2-|\pi_t l|_H^2\right)\theta\circ X\right]\\
&= E_\mu\left[\left((<\pi_t l, V\circ X>_{W^\star,W})^2-|\pi_t l|_H^2\right)\theta\circ X\right]\\
&= E_{X\mu}\left[\left((<\pi_t l, V>_{W^\star,W})^2-|\pi_t l|_H^2\right)\theta\right]\\
&= E_\nu\left[\left((<\pi_t l, V>_{W^\star,W})^2-|\pi_t l|_H^2\right)\theta\right]\\
&= E_\nu\left[\left((<\pi_s l, V>_{W^\star,W})^2-|\pi_s l|_H^2\right)\theta\right]\\
&= E_\mu\left[\left((<\pi_s l, V\circ X>_{W^\star,W})^2-|\pi_s l|_H^2\right)\theta\circ X\right]\\
&= E_\mu\left[\left(\left(<\pi_s l, R^X>_{W^\star,W}\right)^2-|\pi_s l|_H^2\right)\theta\circ X\right]\\
&= E_\mu\left[\left(\left(\delta^{R^X}\pi_s l\right)^2-|\pi_s l|_H^2\right)\theta\circ X\right]
\end{aligned}$$

Where we also used that $(V, H, \pi_.)$ is an abstract Wiener process on $(W, \mathcal{F}_.^\nu, \nu)$ □

Remark VI.2. *When (W, H, μ) is the classical Wiener space and $\pi_.$ the sequence defined by $(\pi_t h)_. = \int_0^{.\wedge t}\dot{h}_s ds$. Let $U = I_W + u$ be such that $U\mu = \nu$, and $R^U := V\circ U$, where $V = I_W + v$ is the Girsanov shift associated with $U\mu$. Then $\left(U, R^U\right)$ is the weak solution of the stochastic differential equation*

$$dU_t = dR_t^U - \dot{v}_t\circ U dt, U_0 = 0$$

on the space $\left(W, \mathcal{F}_.^U, \mu\right)$. In particular it holds when U is the solution to the Monge problem. Hence the Theorem 11 of **[47]** *appears as a straightforward application of this theorem.*

5. First applications, the innovation process

5.1. Adapted perturbations of the identity. We recall here that throughout this paper, for any probability ν which is absolutely continuous with respect to the Wiener measure and for any $u\in L^0(\nu, H)$ we set

$$\rho\left(-\delta^W u\right) := \exp\left(-\delta^W u - \frac{|u|_H^2}{2}\right)$$

Definition VI.8. *Let $\nu << \mu$ be a probability which is absolutely continuous with respect to μ and $u\in L^0(\nu, H)$, and $U = I_W + u$. Then we will say that U is a perturbation of the identity ($\nu - PI$). If u is in $L_a^0(\nu, H)$, we shall say that $U = I_W + u$ is an adapted perturbation of the identity ($\nu - API$). If u is in $L_a^2(\nu, H)$ and $E_\mu\left[\rho\left(-\delta^W u\right)\right] = 1$, we will say that $U = I_W + u$ is a nice adapted perturbation of the identity ($\nu - NAPI$). When $\nu = \mu$ we will say API (resp. PI resp. $NAPI$) instead of $\mu - API$ (resp. $\mu - PI$, $\mu - NAPI$).*

Proposition VI.8. *Let ν be a probability which is absolutely continuous with respect to μ and $V = I_W + v\in \mathcal{R}_a(\nu, \mu)$ be the associated Girsanov shift. Moreover, assume that there is a $U = I_W + u\in\mathcal{R}_a(\mu, \nu)$ with $u\in L_{loc}^2(\mu, H)$ Then we have $\mu - a.s.$*

$$\left(\delta^W v\right)\circ U = \delta^W(v\circ U) + < v\circ U, u>_H$$

Proof: Suppose first that v is of the form $v = \sum_i\left(\pi_{t_{i+1}} - \pi_{t_i}\right)h_i a_i$ where $(h_n)_{n\in\mathbb{N}}$ is an orthonormal basis of H, and $a_i\in L^2\left(\nu, \mathcal{F}_{t_i}^\nu\right)$. As a matter of fact, since $W^\star$ is dense in H we can even restrict ourself to v of the form $v = \sum_{i=1}^n\left(\pi_{t_{i+1}} - \pi_{t_i}\right)l_i a_i$ where for every i $l_i\in W^\star$, and where the (l_i) are orthogonal in H , and a_i as above. Then we have :

$$\begin{aligned}
\left(\delta^W v\right)\circ U &= \left(\delta^V v - |v|_H^2\right)\circ U \\
&= \sum_{i=1}^n \left(a_i\circ U < \left(\pi_{t_{i+1}}-\pi_{t_i}\right) l_i, V\circ U >_{W^\star,W}\right) - |v\circ U|_H^2 \\
&= \sum_{i=1}^n \left(a_i\circ U < \left(\pi_{t_{i+1}}-\pi_{t_i}\right) l_i, I_W + u + v\circ U >_{W^\star,W}\right) - |v\circ U|_H^2 \\
&= \sum_{i=1}^n a_i\circ U\left(\delta^W\pi_{t_{i+1}} l_i - \delta^W \pi_{t_i} l_i\right) + < \sum_{i=1}^n \left(\pi_{t_{i+1}}-\pi_{t_i}\right) l_i a_i\circ U, u + v\circ U >_{W^\star,W} - |v\circ U|_H^2 \\
&= \delta^W\left(v\circ U\right) + < v\circ U, u+v\circ U >_{W^\star,W} - |v\circ U|_H^2 \\
&= \delta^W\left(v\circ U\right) + < v\circ U, u >_H
\end{aligned}$$

Let $(v_n)_{n\in\mathbb{N}} \subset L_a^2(\nu,H)$ be a sequence which converges strongly in $L^2(\nu,H)$ to $v\in L^2(\nu,H)$. Since $U\mu=\nu$ we also have that the sequence $(v_n\circ U)_{n\in\mathbb{N}}\subset L_a^2(\mu,H)$ converges strongly in $L^2(\mu,H)$ to $v\circ U\in L^2(\mu,H)$. Thus $\delta^W(v_n\circ U)\to\delta^W(v\circ U)$ in $L^2(\mu)$. Moreover, since $U\mu=\nu$, the convergence of $\delta^W v_n$ to $\delta^W v$ in $L^2(\nu)$ implies the convergence of $\left(\delta^W v_n\right)\circ U$ to $\left(\delta^W v\right)\circ U$ in $L^2(\mu)$. Hence, the results holds again for $v\in L^2(\nu,H)$ by density. Finally, if $v\in L^2_{loc}(\nu,H)$, let τ_n a sequence which reduces v, we then have for every $n\in\mathbb{N}$

$$\left(\delta^W\pi_{\tau_n} v\right)\circ U = \delta^W\left(\pi_{\tau_n} v\circ U\right) + < \left(\pi_{\tau_n} v\right)\circ U, u >_H$$

Since $U\mu=\nu$, $(\tau_n\circ U)_{n\in\mathbb{N}}$ also reduces $v\circ U$ on $(W,\mathcal{F}^\mu_\cdot,\mu)$ and we can take the limit. □

5.2. Some straightforward corollaries.

Corollary VI.2. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ and such that $d(\mu,\nu)<\infty$. We further note $V=I_W+v\in L_a^0(\nu,W)$ the Girsanov shift associated with ν. Then*

$$d(\mu,\nu)=\min\left(E_\nu\left[|S-V|_H^2\right] : S\in\mathcal{R}(\nu,\nu)\right) \tag{5.57}$$

Moreover the optimal rotation S^T is such that $S^T=T\circ V$ where T is the solution of the Monge problem (see Theorem VI.2).

Proof: By Theorem VI.3 we have $V\nu=\mu$. On the other hand for any $S\in\mathcal{R}(\nu,\nu)$ by definition $S\nu=\nu$. Hence $(V\times S)^*\nu\in\Sigma(\mu,\nu)$ so that

$$d(\mu,\nu)\le E_\nu\left[|S-V|_H^2\right]$$

To see that the optimum is attained, let T be the solution of the Monge problem. Since $T\in\mathcal{R}(\mu,\nu)$ we get

$$\begin{aligned}
d(\mu,\nu) &= E_\mu\left[|T-I_W|_H^2\right] \\
&= E_{V\nu}\left[|T-I_W|_H^2\right] \\
&= E_\nu\left[|T\circ V-V|_H^2\right]
\end{aligned}$$

Together with the fact that $V\in\mathcal{R}(\nu,\mu)$, from the Proposition VI.7 we know that $S^T=T\circ V\in\mathcal{R}(\nu,\nu)$. Thus the infimum of (5.57) is attained by S_T. □

Note that this corollary can be formulated as the research of an optimal element of $\mathcal{R}_a(\nu,\mu)$

Corollary VI.3. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ such that $d(\nu,\mu) < \infty$. Then there is a rotation $S \in \mathcal{R}(\nu,\nu)$ such that*

$$d(\mu,\nu) = \min\left(\left\{E_\nu\left[|S-A|_H^2\right] : A \in \mathcal{R}_a(\nu,\mu)\right\}\right) \tag{5.58}$$

Moreover the infimum is attained by a $V \in \mathcal{R}_a(\nu,\mu)$ such that $S = T \circ V$

Proof: It is the same principle as in the proof of Corollary VI.2. □

The next Corollary has to be compared with the Theorem 1 of [**11**]

Corollary VI.4. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ and $V = I_W + v \in \mathcal{R}_a(\nu,\mu)$ be the Girsanov shift associated with ν. Then for every $X \in \mathcal{R}(\mu,\nu)$ there is a unique $\widetilde{x} \in L^0(\mu,H)$ which is adapted to the augmented filtration $\mathcal{F}^X_.$ generated by X, and a unique $R \in \mathcal{R}(\mu,\mu)$ with the property that $(R,H,\pi_.)$ is an abstract Wiener process on $(W,\mathcal{F}^X_.,\mu)$ such that*

$$X = R + \widetilde{x}$$

Moreover we have $R = V \circ X$ and $\widetilde{x} + v \circ X = 0$.

5.3. The innovation process.

Definition VI.9. *Let $U = I_W + u \in L^0(\mu,W)$ where $u \in L^2(\mu,H)$. The dual predicable projection $\widehat{u}$ of u with respect to the augmented filtration $\mathcal{F}^U_.$ generated by U, is defined as the projection of u on the closed subspace of the elements of $L^2(\mu,H)$ which are adapted to $\mathcal{F}^U_.$. Moreover, the abstract innovation of $U = I_W + u$, $u \in L^2(\mu,H)$, is defined by $Z^U = U - \widehat{u}$ and the innovation process of U is defined by the triplet $\left(Z^U,H,\pi_.\right)$.*

The dual predicable projection $\widehat{u}$ of an adapted u is sometimes called its causal estimate. Its main property is the following which is well known :

Proposition VI.9. *Let $U = I_W + u$ be in $L^2_a(\mu,W)$, with $u \in L^2_a(\mu,H)$, and $\mathcal{F}^U_.$ be the augmented filtration generated by U. Then $\left(Z^U,H,\pi_.\right)$ is an abstract Wiener process on $\left(W,\mathcal{F}^U_.,\mu\right)$ where Z^U is the innovation of U.*

Proof: For every $l \in W^\star$, for every $0 \leq s < t \leq 1$ and for every bounded function θ on W which is $\mathcal{F}^\mu_s$ measurable, we have :

$$E_\mu\left[< (\pi_t - \pi_s)\, l, Z^U >_{W^\star,W} \theta \circ U\right] \quad = \quad A + B$$

where

$$A := E_\mu\left[< (\pi_t - \pi_s)\, l, I_W >_{W^\star,W} \theta \circ U\right]$$

and

$$B := E_\mu\left[< (\pi_t - \pi_s)\, l, u - \widehat{u} >_{W^\star,W} \theta \circ U\right]$$

Since u is adapted $\mathcal{F}^U_s \subset \mathcal{F}^\mu_s$, the first term of the right hand term is null. The second term can be written :

$$E_\mu\left[< \theta \circ U\, (\pi_t - \pi_s)\, l, u - \widehat{u} >_{W^\star,W}\right]$$

Since $\theta \circ U\, (\pi_{t\vee s} - \pi_s)\, l$ is clearly adapted to $\mathcal{F}^U_.$, by definition of the dual predicable projection

$$E_\mu\left[< \theta \circ U\, (\pi_t - \pi_s)\, l, u - \widehat{u} >_{W^\star,W}\right] = 0$$

for every $s < t$. This shows that Z^U is an abstract martingale. Moreover, since $t \to < u - \widehat{u}, \pi_t l >_H$ is of finite variation, $< \delta^{Z^U} \pi_. l >_t = < \delta^W \pi_. l >_t = |\pi_t l|^2_H$. Thus the proof is achieved. □

The following Corollary provides a necessary and sufficient condition for the innovation process to be a Wiener process.

Corollary VI.5. *Let ν be a probability which absolutely continuous with respect to the Wiener measure μ. Let $V = I_W + v \in \mathcal{R}_a(\nu,\mu)$ denote the Girsanov shift of ν. Moreover, assume that there is a $U \in \mathcal{R}(\mu,\nu)$ of the form $U = I_W + u$ where $u \in L^2(\mu,H)$, and let $\mathcal{F}^U_\cdot$ be the augmented filtration generated by U. Then the innovation process of U, noted by Z^U, is a $\left(W,\mathcal{F}^U_\cdot,\mu\right)$-abstract Wiener process if and only if $\mu - a.s.$*

$$V \circ U = Z^U$$

i.e. $\mu - a.s.$

$$\widehat{u} + v \circ U = 0$$

In particular for any $U = I_W + u \in \mathcal{R}_a(\mu,\nu)$ with $u \in L^2_a(\mu,H)$, we have

- *$V \circ U = Z^U$ where Z^U is the innovation of U i.e. $\widehat{u} + v \circ U = 0$*
- *$v \in L^2_a(\nu,H)$ and $v \circ U \in L^2_a(\mu,H)$*

Proof: Assume that $V \circ U = Z^U$. Then from the theorem VI.5 we have $Z^U = R^U$ where $\left(R^U,H,\pi_\cdot\right)$ is a $\left(W,\mathcal{F}^U_\cdot,\mu\right)$ abstract Wiener process. Conversely, assume that $\left(Z^U,H,\pi_\cdot\right)$ is a $\left(W,\mathcal{F}^U_\cdot,\mu\right)$ abstract Wiener process. By linearity of the duality product, for all l in $W^\star$ we get :

$$\delta^{R^U}\pi_\cdot l = \delta^{Z^U}\pi_\cdot l + < \pi_\cdot l, \hat{u} + v \circ U >_H$$

which reads

$$\delta^{R^U}\pi_\cdot l - \delta^{Z^U}\pi_\cdot l = < \pi_\cdot l, \hat{u} + v \circ U >_H$$

Since both $\left(Z^U,H,\pi_\cdot\right)$ and $\left(R^U,H,\pi_\cdot\right)$ are $\left(W,\mathcal{F}^U_\cdot,\mu\right)$ abstract Wiener process, the left hand side of the latter equation is a $\mathcal{F}^U_\cdot$ martingale. Thus for all $l \in W^\star$, the process $(< \pi_\cdot l, \hat{u} + v \circ U >_H)$ vanish as a $\mathcal{F}^U_\cdot$ martingale with finite variation, which means that $\hat{u} + v \circ U = 0$ i.e. $V \circ U = Z^U$. As an application of these result, assume that moreover $U \in \mathcal{R}_a(\mu,\nu)$ and $u \in L^2_a(\mu,H)$. Then, we know from proposition VI.9 that $\left(Z^U,H,\pi_\cdot\right)$ is a $\left(W,\mathcal{F}^U_\cdot,\mu\right)$ abstract Wiener process. Hence we have $\mu - a.s.$ $V \circ U = Z^U$. Finally, since $u \in L^2_a(\mu,H)$ then we have

$$|v \circ U|_{L^2(\mu,H)} = |\hat{u}|_{L^2(\mu,H)} \leq |u|_{L^2(\mu,H)} < \infty$$

□

6. Entropy, energy, and mean square error

The next corrolary shows that the entropy is equal to the energy if and only if the mean square error of the causal estimate of an adapted signal u is null (See **[32]** for analogous relationship between the mutual information and the error). The Theorem VI.6 which will relate this result to the invertibility of the Girsanov shift.

Corollary VI.6. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ. Moreover, assume that there is a $U = I_W + u \in \mathcal{R}_a(\mu,\nu)$ such that $u \in L^2_a(\mu,H)$. Then*

$$2H(\nu|\mu) = |\widehat{u}|^2_{L^2(\mu,H)}$$

where $\widehat{u}$ is the dual predicable projection of u on the augmented filtration $\mathcal{F}^U_\cdot$ generated by U. In particular we also have

$$\widehat{\epsilon} = 2\left(E_\mu\left[\frac{|u|^2_H}{2}\right] - H(\nu|\mu)\right)$$

where $\widehat{\epsilon} = |u - \widehat{u}|^2_{L^2(\mu,H)}$ is the mean square error of the causal estimate of u.

Proof: Let Z^U be the innovation of U i.e. $Z^U := U - \hat{u}$. Since $U \in \mathcal{R}_a(\mu,\nu)$ and $u \in L^2_a(\mu, H)$, we know from Corollary VI.5 that $V \circ U = Z^U$, which also reads $\widehat{u} + v \circ U = 0$. Therefore

$$\begin{aligned} E_\mu\left[|\widehat{u}|^2_H\right] &= E_\mu\left[|v \circ U|^2_H\right] \\ &= E_\nu\left[|v|^2_H\right] \\ &= 2H(\nu|\mu) \end{aligned}$$

where the last equality is due to Proposition VI.5 . Since $\widehat{\epsilon} = |u - \widehat{u}|^2_{L^2(\mu,H)} = |u|^2_{L^2(\mu,H)} - |\widehat{u}|^2_{L^2(\mu,H)}$ we get the last equality. □

7. Variational formulation of the entropy

Lemma VI.2. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ. Let $U \in \mathcal{R}(\mu,\nu)$ and $V \in \mathcal{R}(\nu,\mu)$. Then V is $\nu - as$ injective with a left inverse U, if and only if it is $\nu - as$ surjective with a right inverse U. Moreover, in that case both U are the same and V is $\nu-$invertible with inverse U. Furthermore we have the following criterion. Let $V \in \mathcal{R}(\nu,\mu)$, and assume that there is a $U \in L^0(\mu, W)$ which is such that $U\mu << \nu$ and $\mu - a.s.$ $V \circ U = I_W$. Then $U \circ V = I_W$ $\nu - a.s.$ if and only if $U\mu \sim \nu$. Moreover, in that case we also have $U \in \mathcal{R}(\mu,\nu)$ and V is invertible with inverse U.*

Proof: We first prove the second part of the claim. Let $V \in \mathcal{R}(\nu,\mu)$ and $U \in L^0(\mu, W)$ be such that $U\mu \sim \nu$ and $\mu - a.s.$ $V \circ U = I_W$. Then we have :

$$\begin{aligned} U\mu\left(\{\omega | U \circ V = I_W\}\right) &= \mu\left(U^{-1}\left(\{\omega | U \circ V = I_W\}\right)\right) \\ &= \mu\left(\{\omega | U \circ V \circ U = U\}\right) \\ &\geq \mu\left(\{\omega | V \circ U = I_W\}\right) \\ &= 1 \end{aligned}$$

where we have used that $\{\omega | V \circ U = I_W\} \subset \{\omega | U \circ V \circ U = U\}$. Since $U\mu \sim \nu$ we also have $\nu - a.s.$ $U \circ V = I_W$. In particular

$$\begin{aligned} E_\mu[f \circ U] &= E_\nu[f \circ U \circ V] \\ &= E_\nu[f] \end{aligned}$$

so that $U \in \mathcal{R}(\mu,\nu)$. Conversely if $V \in \mathcal{R}(\nu,\mu)$, and $\nu - a.s.$ $U \circ V = I_W$ for a U such that $U\mu << \nu$, then we have $U\mu = \nu$, and in particular $U\mu \sim \nu$.

We now prove the first part of the claim. Assume that $V \in \mathcal{R}(\nu,\mu)$ is μ almost surely surjective with a right inverse $U \in \mathcal{R}(\mu,\nu)$ i.e. $\mu\left(\{\omega | V \circ U = I_W\}\right))$. Since $U \in \mathcal{R}(\mu,\nu)$ we have in particular that $U\mu \sim \nu$ so that $\nu - a.s.$ $U \circ V = I_W$. Conversely, assume that $V \in \mathcal{R}(\nu,\mu)$ is $\nu - as$ injective with a left inverse $U \in \mathcal{R}(\mu,\nu)$, i.e. $\nu\left(\{\omega | U \circ V = I_W\}\right)$. We then have :

$$\begin{aligned} \mu\left(\{\omega | V \circ U = I_W\}\right) &= V\nu\left(\{\omega | V \circ U = I_W\}\right) \\ &= \nu\left(V^{-1}\left(\{\omega | V \circ U = I_W\}\right)\right) \\ &= \nu\left(\{\omega | V \circ U \circ V = V\}\right) \\ &\geq \nu\left(\{\omega | U \circ V = I_W\}\right) \\ &= 1 \end{aligned}$$

Where we have used that $\{\omega | U \circ V = I_W\} \subset \{\omega | V \circ U \circ V = V\}$. Hence we have $\mu - a.s.$ $V \circ U = I_W$. □

Remark: Let $U \in \mathcal{R}_a(\mu,\nu)$ where $\nu << \mu$ is a probability, and let $V = I_W + v \in \mathcal{R}_a(\nu,\mu)$ be the Girsanov shift associated with ν. Further, assume that there is a $\widetilde{V} \in \mathcal{R}(\nu,\mu)$ such that $\nu - a.s.$ $U \circ \widetilde{V} = I_W$.

Then we have not necessarily $\nu - a.s.$ $V = \tilde{V}$. As a matter of fact if we set $R^U := V \circ U$ (which we know to be a rotation in $\mathcal{R}(\mu,\mu)$ and even an abstract Wiener process), then we have $\nu - a.s.$ $V = V \circ \left(U \circ \widetilde{V}\right) = (V \circ U) \circ \widetilde{V} = R^U \circ \widetilde{V}$. In the same way, assume that there is a $\widetilde{V} \in \mathcal{R}(\nu,\mu)$ such that $\mu - a.s.$ $\widetilde{V} \circ U = I_W$. If we set $S^U = U \circ V$ (which we know to be a rotation in $\mathcal{R}(\nu,\nu)$), then we have $\nu - a.s.$ $V = \left(\widetilde{V} \circ U\right) \circ V = \widetilde{V} \circ (U \circ V) = \widetilde{V} \circ S^U$.

The next definition formalize the notions of invertibility of Lemma VI.2

Definition VI.10. *Let ν be a probability such that $\nu << \mu$ (i.e. absolutely continuous). Let $U \in \mathcal{R}(\mu,\nu)$ and $V \in \mathcal{R}(\nu,\mu)$. Then V is said to be $\nu -$ invertible with an inverse U if and only if $\nu - a.s.$ $U \circ V = I_W$ and $\mu - a.s.$ $V \circ U = I_W$.*

Note that when $\nu \sim \mu$, we recover the definition of **[47]**, and we just say that V is invertible.

Proposition VI.10. *Let ν be a probability such that $\nu << \mu$, and let $V \in \mathcal{R}_a(\nu,\mu)$ be the associated Girsanov shift. Further assume that there is a $U \in L^0(\mu, W)$ such that $V \circ U = I_W$ μ as and $U\mu << \nu$, then we have :*

$$\mathcal{F}^{\mu}_{.} \subset \mathcal{F}^{U}_{.}$$

where $\mathcal{F}^{U}_{.}$ is the augmented filtration associated with U. In particular, if U is adapted we have $\mathcal{F}^{\mu}_{.} = \mathcal{F}^{U}_{.}$

Proof: For every $l \in W^\star$ and $A \in \mathcal{B}(\mathbb{R})$, and $t \in [0,1]$, we note $l_t = \pi_t l$ and $\Omega = \{\omega | V \circ U = I_W\}$ which is measurable. We decompose $l_t^{-1}(A)$ in the following way :

$$l_t^{-1}(A) \quad = \quad \left(l_t^{-1}(A) \cap \Omega\right) \cup \left(l_t^{-1}(A) \cap \Omega^c\right)$$

Considering the first term of the right hand side, we have

$$\begin{aligned} l_t^{-1}(A) \cap \Omega \subset U^{-1}(\{\omega | l_t \circ V \in A\}) \subset U^{-1}\left(V^{-1}(\{\omega | l_t \in A\})\right) &= (U)^{-1}(\{\omega | l_t \circ V \in A\}) \\ &= (U)^{-1}\left(\mathcal{F}^{V,0}_t\right) \subset U^{-1}(\mathcal{F}^{\nu}_t) \end{aligned}$$

where we used that V is $\nu - adapted$. Moreover we have

$$U^{-1}(\mathcal{F}^{\nu}_t) = \sigma\left(U^{-1}\left(\mathcal{F}^{W,O}_t\right) \cup U^{-1}(\Theta^{\nu})\right)$$

Since $U\mu << \nu$, $U^{-1}(\Theta^\nu) \subset \Theta^\mu$. Furthermore $U^{-1}\left(\mathcal{F}^{W,O}_t\right) = \mathcal{F}^{U,O}$ Hence $l_t^{-1}(A) \cap \Omega \subset \mathcal{F}^U_t$ On the other hand, $l_t^{-1}(A) \cap \Omega^c$ is in Θ^μ. Finally we have $l_t^{-1}(A) \in \mathcal{F}^U_t$, so that $\mathcal{F}^\mu_t \subset \mathcal{F}^U_t$. The rest of the proof is straightforward. □

Proposition VI.11. *Let ν be a probability such that $\nu << \mu$ (i.e. absolutely continuous with respect to the Wiener measure μ) and $V = I_W + v \in \mathcal{R}_a(\nu,\mu)$ denote the Girsanov shift associated with ν. Moreover, assume that there is a $U = I_W + u \in \mathcal{R}_a(\mu,\nu)$ such that $\mu - a.s.$ $V \circ U = I_W$. Then we also have :*

- $E_\mu\left[\rho\left(-\delta^W u\right)\right] = \mu\left(\frac{d\nu}{d\mu} > 0\right)$
- *V is $\nu-$invertible with inverse U (see Definition VI.10)*
- $\mathcal{F}^{\mu}_{.} = \mathcal{F}^{U}_{.}$ *and* $\mathcal{F}^{\nu}_{.} = \mathcal{F}^{V}_{.}$

Furthermore, if $\mu \sim \nu$ we have $V\mu = \rho\left(-\delta^W u\right)\mu$

Proof: Since both $U \in \mathcal{R}_a(\mu,\nu)$ and $V \in \mathcal{R}_a(\mu,\nu)$, we know from the Proposition VI.8 that

$$\left(\delta^W v\right) \circ U = \delta^W (v \circ U) + < u, v \circ U >_H$$

so that

$$\begin{aligned}
\rho\left(-\delta^W u\right)\rho\left(-\delta^W v\right)\circ U &= \exp\left(-\delta^W u - \frac{|u|_H^2}{2}\right)\times\exp\left(\left(-\delta^W v\right)\circ U - \frac{|v\circ U|_H^2}{2}\right)\\
&= \exp\left(-\delta^W u - \frac{|u|_H^2}{2}\right)\times\exp\left(-\delta^W (v\circ U) - < u, v\circ U >_H - \frac{|v\circ U|_H^2}{2}\right)\\
&= \exp\left(-\delta^W (u + v\circ U) - \frac{|u|_H^2}{2} - < u, v\circ U >_H - \frac{|v\circ U|_H^2}{2}\right)\\
&= \exp\left(-\delta^W (u + v\circ U) - \frac{|u + v\circ U|_H^2}{2}\right)\\
&= \exp\left(-\delta^W (V\circ U - I_W) - \frac{|V\circ U - I_W|_H^2}{2}\right)\\
&= \rho\left(-\delta^W (V\circ U - I_W)\right)
\end{aligned}$$

where we used that $\mu - a.s.$, $V\circ U - I_W = U + v\circ U - I_W = I_W + u + v\circ U - I_W = u + v\circ U$

$$\rho\left(-\delta^W u\right)\rho\left(-\delta^W v\right)\circ U = \rho\left(-\delta^W (V\circ U - I_W)\right)$$

Thus, from the hypothesis $V\circ U = I_W$ $\mu - as$, we get that

$$\rho\left(-\delta^W u\right)\frac{d\nu}{d\mu}\circ U = 1$$

Furthermore $U\mu = \nu$, so that

$$\begin{aligned}
\mu\left(\left\{\omega\middle|\frac{d\nu}{d\mu}\circ U > 0\right\}\right) &= \mu\left(U^{-1}\left(\left\{\omega\middle|\frac{d\nu}{d\mu} > 0\right\}\right)\right)\\
&= \nu\left(\left(\left\{\omega\middle|\frac{d\nu}{d\mu} > 0\right\}\right)\right)\\
&= E_\nu\left[1_{\frac{d\nu}{d\mu}>0}\right]\\
&= E_\mu\left[\frac{d\nu}{d\mu}1_{\frac{d\nu}{d\mu}>0}\right]\\
&= E_\mu\left[\frac{d\nu}{d\mu}\right]\\
&= E_\nu[1]\\
&= 1
\end{aligned}$$

and

$$\begin{aligned}
E_\mu\left[\rho\left(-\delta^W u\right)\right] &= E_\mu\left[\frac{1}{\frac{d\nu}{d\mu}\circ U}\right]\\
&= E_\mu\left[\frac{1}{\frac{d\nu}{d\mu}\circ U}1_{\frac{d\nu}{d\mu}\circ U>0}\right]\\
&= E_\nu\left[\frac{1}{\frac{d\nu}{d\mu}}1_{\frac{d\nu}{d\mu}>0}\right]\\
&= E_\mu\left[\frac{d\nu}{d\mu}\frac{1}{\frac{d\nu}{d\mu}}1_{\frac{d\nu}{d\mu}>0}\right]\\
&= E_\mu\left[1_{\frac{d\nu}{d\mu}>0}\right]\\
&= \mu\left(\omega\middle|\frac{d\nu}{d\mu} > 0\right)
\end{aligned}$$

By applying Proposition VI.10 we have $\mathcal{F}_{\cdot} = \mathcal{F}_{\cdot}^U$. Since from Lemma VI.2 $U \circ V = I_W$ $\nu - a.s.$, the augmented filtration associated with $U \circ V$ is equal to $\mathcal{F}_t^\nu$. Up to a $\nu-$negligible set, we have :

$$\begin{aligned} \sigma\left(< \pi_t l, U \circ V >_{W^\star, W}, l \in W^\star\right) &= V^{-1}\left(\sigma\left(< \pi_t l, U >_{W^\star, W}, l \in W^\star\right)\right) \\ &\subset V^{-1}\left(\mathcal{F}_t^U\right) \\ &= V^{-1}\left(\mathcal{F}_t^\mu\right) \subset \mathcal{F}_t^V \end{aligned}$$

so that $\mathcal{F}_t^\nu \subset \mathcal{F}_t^V$. Since V is adapted to $\mathcal{F}_{\cdot}^\nu$ we get the equality of the filtrations. □

As a consequence we have:

Theorem VI.6. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ and let $V = I_W + v$ denote the Girsanov shift of ν. Then, for every $U = I_W + u \in \mathcal{R}_a(\mu, \nu)$ the following inequality holds*

$$2H(\nu|\mu) \leq E_\mu\left[|u|_H^2\right]$$

Moreover, if $H(\nu|\mu) < \infty$ we have the following equations

$$2H(\nu|\mu) = E_\mu\left[|u|_H^2\right] - E_\mu\left[|V \circ U - I_W|_H^2\right] \tag{7.59}$$

and

$$\widehat{\epsilon} = E_\mu\left[|V \circ U - I_W|_H^2\right]$$

where $\widehat{\epsilon}$ denotes the mean square error of the causal estimate of u which is defined by $\widehat{\epsilon} := E_\mu\left[|u - \widehat{u}|_H^2\right]$

In particular, in that case the following assertions are equivalent :

- $2H(\nu|\mu) = E_\mu\left[|u|_H^2\right]$
- $\widehat{\epsilon} = 0$
- *V is $\nu-$invertible with inverse U (see Definition VI.10)*

and in that case $E_\mu[\rho(-\delta u)] = \mu\left(\left\{\omega \in W : \frac{d\nu}{d\mu} > 0\right\}\right)$

Proof: Assume that $u = U - I_W$ is not in $L^2(\mu, H)$. Then the inequality is true since $|u|_{L^2(\mu,H)} = \infty$. Now suppose that $u \in L_a^2(\mu, H)$. Since both $U \in \mathcal{R}_a(\mu, \nu)$ and $V \in \mathcal{R}_a(\nu, \mu)$, we have (see the proof of Proposition VI.11) :

$$\rho\left(-\delta^W u\right)\rho\left(-\delta^W v\right) \circ U = \rho\left(-\delta^W(V \circ U - I_W)\right)$$

Therefore

$$\begin{aligned} H(\nu|\mu) &= E_\mu\left[\ln \rho\left(-\delta^W v\right) \circ U\right] \\ &= E_\mu\left[\ln \rho\left(-\delta^W(V \circ U - I_W)\right)\right] - E_\mu\left[\ln \rho\left(-\delta^W u\right)\right] \\ &= E_\mu\left[-\delta^W(V \circ U - I_W) - \frac{1}{2}|V \circ U - I_W|_H^2\right] + E_\mu\left[\delta u + \frac{|u|_H^2}{2}\right] \end{aligned}$$

From the last point of Corollary VI.5 we know that $v \circ U \in L_a^2(\mu, H)$, so that $(V \circ U - I_W) = u + v \circ U \in L_a^2(\mu, H)$. Thus both $\delta^W \pi_{\cdot}(V \circ U - I_W)$ and $\delta^W \pi_{\cdot}(u)$ are martingale vanishing at 0, and $E_\mu\left[\delta^W u\right] = E_\mu\left[\delta^W(V \circ U - I_W)\right] = 0$. Hence we get :

$$2H(\nu|\mu) = E_\mu\left[|u|_H^2\right] - E_\mu\left[|V \circ U - I_W|_H^2\right] \tag{7.60}$$

which yields the inequality in the case $u \in L^2(\mu, H)$. Thus, we have proved that the inequality holds for all adapted measurable mapping $U \in \mathcal{R}_a(\mu, \nu)$ which represents ν. Assume that $H(\nu|\mu) < \infty$ and that the equality occurs in equation 7.59 for an API U such that $U\mu = \nu$. Since $u \in L^2(\mu, H)$ we have $V \circ U = I_W$

i.e. U is the right inverse of V. Since U is an API which represents ν, the Proposition VI.11 shows that U is the inverse of V, and

$$E_\mu\left[\rho\left(-\delta u\right)\right] = \mu\left(\left\{\omega \in W : \frac{d\nu}{d\mu} > 0\right\}\right)$$

Conversely, assume that V is the almost sure inverse of U and that the entropy is finite. Then we have:

$$E_\mu\left[|u|_H^2\right] = E_\mu\left[|v \circ U|_H^2\right] = E_\nu\left[|v|_H^2\right] = H\left(\nu|\mu\right)$$

where the last equality comes from Proposition VI.5. To end the proof with, note that since from Corollary VI.5 we have $\widehat{u} + v \circ U = 0$, we also have $E_\mu\left[|V \circ U - I_W|_H^2\right] = E_\mu\left[|u + v \circ U|_H^2\right] = E_\mu\left[|u - \widehat{u}|_H^2\right] = \widehat{\epsilon}$ □

Remark VI.3. *Although we could have deduced this theorem from Corollary VI.6 and from the condition $\widehat{u} + v \circ U = 0$, we preferred to give a direct proof.*

An alternative way to formulate the latter theorem is the following Corollary VI.7 :

Corollary VI.7. *Let ν be a probability with finite entropy $H\left(\nu|\mu\right) < \infty$ which implies $\nu << \mu$, and let $V = I_W + v \in \mathcal{R}_a\left(\nu, \mu\right)$ denote the Girsanov shift associated with ν. Then the following assertions are equivalent :*

- *The following variational representation of the entropy holds :*

$$2H\left(\nu|\mu\right) = \min\left(E_\mu\left[|u|_H^2\right] : U = I_W + u \in \mathcal{R}_a\left(\mu, \nu\right)\right) \tag{7.61}$$

- *V is $\nu-$invertible (see Definition VI.10) i.e. there is a $U = I_W + u \in \mathcal{R}_a\left(\mu, \nu\right)$, where $u \in L_a^2\left(\mu, H\right)$ which is the almost sure inverse of $V = I_W + v$*

When it occurs, let U denote the inverse of V, then the infimum of the variational formula is attained by U and $E_\mu\left[\rho\left(-\delta u\right)\right] = \mu\left(\left\{\omega \in W : \frac{d\nu}{d\mu} > 0\right\}\right)$.

Proof: Assume that there is an API $U = I_W + u \in \mathcal{R}_a\left(\mu, \nu\right)$, where $u \in L_a^0\left(\mu, H\right)$ which is the inverse of V, i.e. $\mu - a.s.$ $V \circ U = I_W$ and $\nu - a.s.$ $U \circ V = I_W$. Since $V \circ U = I_W$ and the entropy is finite, from the Theorem VI.6 we get $2H\left(\nu|\mu\right) = |u|^2_{L^2(\mu,H)}$. Let $A = I_W + a$ be another API in $\mathcal{R}_a\left(\mu, \nu\right)$, then by applying the Theorem VI.6 we get that $H\left(\nu|\mu\right) \leq E_\mu\left[|u|_H^2\right]$. Therefore we get the formula 7.61, with an optimum attained in U. Conversely, assume that we have formula 7.61, and let U be the optimal shift. Since the entropy is finite, u is in $L^2\left(\mu, H\right)$ and by Theorem VI.6 we know that V is $\nu-$invertible with inverse U □

Remark VI.4. *Of course, contrary to equation 7.61, the next (weaker) formula (of stochastic control) holds for any $\nu << \mu$:*

(7.62)

$$2H\left(\nu|\mu\right) = \min\left(\left\{E_{\widetilde{\nu}}\left[|U - B|_H^2\right] : U \in \mathcal{R}_a\left(\widetilde{\nu}, \nu\right), B \in \mathcal{R}_a\left(\widetilde{\nu}, \mu\right) \text{ and } B\mathcal{F}^{\widetilde{\nu}} - \text{Brownian}, \widetilde{\nu} \text{ probability }, \widetilde{\nu} << \mu\right\}\right)$$

where $\mathcal{R}_a\left(\widetilde{\nu}, \nu\right)$ denotes the subset of the $U \in \mathcal{R}\left(\widetilde{\nu}, \nu\right)$ for which there is a $u \in L_a^0\left(\widehat{\nu}, H\right)$ such that $\widehat{\nu} - a.s.$ $U = B + u$. Indeed the optimum is always attained by $(U, B) = (I_W, V)$ on $(W, \mathcal{F}^\nu, \nu)$ t(i.e. for $\widetilde{\nu} = \nu$), and in that case the inequality can be proved with analogous calculus as for equation 7.61 .

8. Local properties

In this section we check that the general properties we proved in part I still hold for the $\nu-$invertibility. Thanks to these extensions, the applications of Part I to statistical mechanics and to information theory easily extend to the case of probabilities absolutely continuous with respect to the Wiener measure (but not necessarily equivalent).

Proposition VI.12. *Let ν be a probability absolutely continuous with respect to μ whose Girsanov shift is noted $V := I_W + v$. Further assume that V is $\nu-$invertible with an inverse $U := I_W + u$. Then, for any $(\mathcal{F}_t^\nu)$ optional time τ, $V^\tau := I_W + \pi_\tau v$ is ν^τ invertible with inverse*

$$U^\tau := I_W + \pi_{\tau\circ U} u$$

where ν^τ is the probability absolutely continuous with respect to ν whose density is defined by $\nu - a.s.$ $\frac{d\nu^\tau}{d\nu} := \exp\left(\delta^V\left(\left(I_H - \pi_\tau\right)v\right) - \frac{|(I_H - \pi_\tau)v|_H^2}{2}\right)$

Proof: First note that the Girsanov theorem directly yields that V^τ is the Girsanov shift associated with ν^τ and that since $U\mu << \nu$ we have $U\mu << \nu^\tau$. Thus, the proof can be obtained in the same way as in Proposition II.2 : we adopt here the same notations, and we follow its steps. First note that the Theorem 59 of **[6]** still applies by using ν instead of μ, and that since τ is defined νa.s. and $U\mu = \nu$, the inverse images are still well defined. By working on simple processes (we do it in **[27]**), after a standard limiting procedure, it is staightforward to check that $\pi_{T\circ U}\left(v \circ U\right) = \pi_{T\circ U}\left(v \circ U^\pi\right)$. On the other hand, by definition $< \pi_s l, U\left(\omega\right) >=< \pi_s l, U^\pi\left(\omega\right) >$ for any $l \in W^\star$ and $s \leq T\left(U\left(\omega\right)\right)$. Moreover the proof of Theorem 100.a of **[6]** clearly still holds mutatis mutandis (just change $W_t\left(\omega\right) = W_t\left(\widetilde{\omega}\right)$ by $< \pi_t l, \omega >=< \pi_t l, \widetilde{\omega} >$ for any $l \in W^\star$ and follow the steps of the proof). Finally, the end of the proof is the same as Proposition II.2. □

Definition VI.11. *Let ν be a probability absolutely continuous with respect to the Wiener measure whose Girsanov shift is noted $V := I_W + v$. We say that V is locally invertible if there is an almost surely increasing sequence of $(\mathcal{F}_t^\nu)-$optional times (τ_n) such that $\nu - a.s.$*

$$\tau_n \uparrow 1$$

and for any $n \in \mathbb{N}$ $V^n := I_W + \pi_{\tau_n} v$ is ν^n invertible, where $\nu - a.s.$

$$\frac{d\nu^n}{d\nu} := \exp\left(\delta^V\left(\left(I_H - \pi_{\tau_n}\right)v\right) - \frac{|\left(I_H - \pi_{\tau_n}\right)v|_H^2}{2}\right)$$

and where I_H is the identity map on H

Theorem VI.7. *Let ν be a probability absolutely continuous with respect to the Wiener measure μ. Then V is locally $\nu-$invertible if and only if it is $\nu-$invertible*

Proof: Since $\nu << \mu$, Lusin's theorem still applies as in the proof of Theorem III.2 and we only have to check that the inverse images are well defined. As a matter of fact it is sufficient to check that for any $n \geq m$ the inverse image $(\pi_{\tau_n} v) \circ U^m$ is well defined where U^m is the ν^{τ_m} inverse of $V^m := I_W + \pi_{\tau_m} v$ and where $\nu^{\tau_m} := \exp\left(\delta^V\left(\left(I_H - \pi_{\tau_m}\right)v\right) - \frac{|\left(I_H - \pi_{\tau_m}\right)v|_H^2}{2}\right).\nu$. Since $U^m\mu = \nu^{\tau_m}$, we just have to check that $\nu^{\tau_m} << \nu^{\tau_n}$ whenever $n \geq m$, which is the case (we even have

$$\nu^m := \exp\left(\delta^{V^n}\left(\left(\pi_{\tau_n} v - \pi_{\tau_m}\right)v\right) - \frac{|\left(\pi_{\tau_n} - \pi_{\tau_m}\right)v|_H^2}{2}\right).\nu^n$$

where $V^n := I_W + \pi_{\tau_n} v$). □

9. The Monge problem and the invertibility

The following lemma gives a polar decomposition of the $\mathcal{R}(\mu,\nu)$ under an invertibility assumption.

Lemma VI.3. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ and let $V = I_W + v \in \mathcal{R}_a(\nu,\mu)$ denote the associated Girsanov shift. Further, assume that V is $\nu-$ invertible (see Definition VI.10) with an inverse $U \in \mathcal{R}(\mu,\nu)$. Then for any $X \in L^0(\mu,W)$, the following assertions are equivalents :*

- $X \in \mathcal{R}(\mu,\nu)$
- $X = U \circ R$ μ *a.s. for a* $R \in \mathcal{R}(\mu,\mu)$

Moreover, in that case we have $R = R^X$ where $R^X := V \circ X$

Proof:

Assume that there is a rotation $R \in \mathcal{R}(\mu,\mu)$ such that $X = U \circ R$. Since $U \in \mathcal{R}(\mu,\nu)$, we have

$$X\mu = (U \circ R)\mu = U(R\mu) = U\mu = \nu$$

In other words X represents ν and $X \in \mathcal{R}(\mu,\nu)$. Conversely, assume that X represents ν. Then, if we set $R^X = V \circ X$, we know by Theorem VI.5 that $R^X \in \mathcal{R}(\mu,\mu)$. Together with the hypothesis that $U \circ V = I_W$ $\nu - a.s.$ this implies

$$X = (U \circ V) \circ X = U \circ (V \circ X) = U \circ R^X$$

Hence X is of the form $X = U \circ R$ where $R = R^X$ is a rotation and even an abstract Wiener process. $\square$

The next result gives the link between the Monge problem and the invertibility of API. It relates the invertibility of the Girsanov shift associated with a law, to the existence of a dual formula to the one given by Corollary VI.2

Theorem VI.8. *Let ν be a probability which is absolutely continuous with respect to the Wiener measure μ, with the property that $H(\nu|\mu) < \infty$. Furthermore let $V = I_W + v \in \mathcal{R}_a(\nu,\mu)$ be the Girsanov shift associated with ν. Then the following assertions are equivalent*

(1) There is an API $U = I_W + u \in \mathcal{R}_a(\mu,\nu)$, $u \in L^2_a(\mu,H)$ such that

$$d(\mu,\nu) = \min\left(E_\mu\left[|U \circ R - I_W|_H^2\right] : R \in \mathcal{R}(\mu,\mu)\right) \tag{9.63}$$

and the infimum is attained for a given R such that $(R,H,\pi_\cdot)$ is an abstract Wiener process on $(W, \mathcal{F}^R_\cdot, \mu)$

(2) V is $\nu-$invertible (see Definition VI.10) with an inverse $U = I_W + u \in \mathcal{R}_a(\mu,\nu)$, $u \in L^2_a(\mu,H)$

In that case all the U are the same and of course U is a $NAPI$. Moreover let T be the optimal solution to the Monge problem given by Theorem VI.2 and let R be the optimal rotation that attains the infimum. Then the polar decomposition of T is given by $T = U \circ R$. In other words, let $R^T := V \circ T$. Then $R^T = R$

Proof: Assume that the first point holds and denote by T shift which is solution of the Monge Problem given in Theorem VI.2, and by R the optimal rotation of the variational problem 9.63. Then, the uniqueness of the solution to the Monge problem implies $T = U \circ R$. Since U is adapted to the filtration generated by the coordinate process $\mathcal{F}^\mu_\cdot$, we have $\mathcal{F}^T_\cdot \subset \mathcal{F}^R_\cdot$. If we set $R^T := V \circ T$, we know from Theorem VI.5 that $\left(R^T, H, \pi_\cdot\right)$ is a $\left(W, \mathcal{F}^T_\cdot, \mu\right)$ abstract Wiener process. Since $(R,H,\pi_\cdot)$ is $\left(W, \mathcal{F}^R_\cdot, \mu\right)$ abstract Wiener process, and $\mathcal{F}^T_\cdot \subset \mathcal{F}^R_\cdot$, this latter is also a $\left(W, \mathcal{F}^T_\cdot, \mu\right)$ abstract Wiener process. Hence, by linearity, $\left(R^T - R, H, \pi_\cdot\right)$ is also a $\left(W, \mathcal{F}^T_\cdot, \mu\right)$ abstract Wiener process. Since $U = I_W + u$ represents ν and is adapted, we know from Corollary VI.5 that $V \circ U = Z^U$ where Z^U is the innovation of U which is defined by $Z^U = U - \widehat{u}$ where

$\widehat{u}$ denotes the dual predicable projection of u with respect to the augmented filtration $\mathcal{F}^U$ generated by U. Then we have

$$R^T = V \circ T = V \circ U \circ R = Z^u \circ R = (U - \hat{u}) \circ R = R + (u - \hat{u}) \circ R$$

Let l be in $W^\star$, then we have $\delta^{R^T-R}\pi_. l =< \pi_. l, (u - \hat{u}) \circ R >_{W^\star,W}$. Since $\left(R^T - R, H, \pi_.\right)$ is a $\left(W, \mathcal{F}^T_., \mu\right)$ abstract Wiener process, the right term of the last equality is a martingale, whereas the left hand term is of finite variation. Therefore the process $\delta^{R^T-R}\pi_. l$ vanish as a martingale of finite variation and we get $R = R^T$ (or equivalently $T = U \circ V \circ T$). On the other hand T is $\nu-$invertible with an inverse which we note K. Thus $\nu - a.s.\ U \circ V = U \circ V \circ (T \circ K) = (U \circ V \circ T) \circ K = T \circ K = I_W$. Then, the Lemma VI.2 implies that V is $\nu-$invertible with inverse U. Therefore we have proved that the first point implies the invertibility of U. Conversely, assume that $V \in \mathcal{R}_a(\nu, \mu)$ is $\nu-$invertible with inverse $U \in \mathcal{R}(\mu, \nu)$ and let $R \in \mathcal{R}(\mu, \mu)$. Then we have $U \circ R \in \mathcal{R}(\mu, \nu)$, so that $(I_W \times U \circ R)\mu \in \Sigma(\mu, \nu)$. Thus the optimality condition reads :

$$d(\mu, \nu) \leq E_\mu\left[|U \circ R - I_W|_H^2\right]$$

Let T be the optimal solution of the Monge problem. Since U is the inverse of V and $T \in \mathcal{R}(\mu, \nu)$, thanks to the lemma VI.3, we know that $T = U \circ R^T$ so that

$$\begin{aligned} d(\mu, \nu) &= E_\mu\left[|T - W|_H^2\right] \\ &= E_\mu\left[|U \circ R^T - W|_H^2\right] \end{aligned}$$

Which shows that the infimum is attained ☐

We already said that the formula 5.57 was equivalent to an optimal problem in terms of elements of $\mathcal{R}_a(\nu, \mu)$. As a matter of fact, the same approach can be performed for formula 9.63. Whenever V is $\nu-$invertible, the following Corollary shows that under suitable conditions, the Monge problem can be seen as the search of an optimal API, but this times in $\mathcal{R}_a(\mu, \nu)$. Thus the next Corollary VI.8 has to be compared with Corollary VI.7 which deals with entropy instead of the Wasserstein distance.

Corollary VI.8. *Let ν be a probability with finite entropy with respect to the Wiener measure μ ($H(\nu|\mu) < \infty$) and let $V = I_W + v \in \mathcal{R}_a(\nu, \mu)$ be the associated Girsanov shift. Then the following assertions are equivalent*

(1) There is an abstract process $(R, H, \pi_.)$ which is an abstract Wiener process on $(W, \mathcal{F}^R_., \mu)$ such that

$$d(\mu, \nu) = \inf\left(\left\{E_\mu\left[|A \circ R - I_W|_H^2\right] : A \in \mathcal{R}_a(\mu, \nu)\right\}\right) \tag{9.64}$$

Moreover the infimum is attained for a given $U = I_W + u \in \mathcal{R}_a(\mu, \nu)$, $u \in L_a^2(\mu, H)$

(2) $V = I_W + v$ is $\nu-$invertible (see Definition VI.10) with an inverse $U = I_W + u \in \mathcal{R}_a(\mu, \nu)$, $u \in L_a^2(\mu, H)$

In that case, all the optimums U are the same, and the optimal rotation of 9.63 is the same that the one noted by R in 9.64. Moreover let T be the optimal solution to the Monge problem, and let R be the optimal rotation that realise the infimum. Then the polar decomposition of T is given by $T = U \circ R$. In other words, $R^T = R$

Proof: The proof follows easily from the equivalence of the $\nu-$invertibility of V and formula 9.63, and by the uniqueness of the optimal solution to the Monge problem just as in the second part of the last proof ☐

10. Application to the innovation conjecture

In this section we use the Theorem VI.6 to enlighten the innovation conjecture : our results extend those of **[47]**. The problem is sumed-up in the next definition, and we refer to **[1]**, **[60]**, or **[47]** for further details on this topic.

Definition VI.12. *Let $A = I_W + a$ where $a \in L^2(\mu, H)$, and Z^A be its innovation. The innovation conjecture is said to be satisfied iff $\mathcal{F}^A = \mathcal{F}^{Z^A}$, where $\mathcal{F}^A$ (resp. $\mathcal{F}^{Z^A}$) is the augmented filtration of A (resp. of Z^A).*

As a straightforward consequence of Lemma VI.3 and Corollary VI.5 we have

Lemma VI.4. *Let ν be a probability such that $H(\nu|\mu) < \infty$ and let $V = I_W + v \in \mathcal{R}_a(\nu, \mu)$ denote the Girsanov shift associated with ν. Moreover, assume that $V = I_W + v \in \mathcal{R}_a(\nu, \mu)$ is $\nu-$ invertible with an inverse $U = I_W + u \in \mathcal{R}(\mu, \nu)$. Finally, let $A = I_W + a$, $a \in L^2(\mu, H)$ be such that its innovation process is an abstract Wiener process on $(W, \mathcal{F}^\mu_\cdot, \mu)$, which happens in particular when $a \in L^2_a(\mu, H)$. Then the following assertions are equivalents :*

- $A\mu = \nu$
- $A = U \circ Z^A$

where Z^A is the innovation process of A

The following theorem shows that, under suitable conditions, the validity of the innovation conjecture only depends on the law of the process, specifically of the invertibility of the Girsanov shift V associated with the law. In particular, note that under suitable conditions the Theorem VI.9 provides a criterion of invertibility for the Girsanov shift associated with a given law.

Theorem VI.9. *Let ν be a probability such that $H(\nu|\mu) < \infty$, and let $V = I_W + v \in \mathcal{R}_a(\nu, \mu)$ be the associated Girsanov shift. Then the following alternative holds:*

- *If V is $\nu-$ invertible with an inverse $U \in \mathcal{R}_a(\mu, \nu)$, then for every $A = I_W + a \in \mathcal{R}_a(\mu, \nu)$, $a \in L^2_a(\mu, H)$, the innovation conjecture is true*
- *If V is $\nu-$invertible with an inverse $U \in \mathcal{R}_a(\mu, \nu)$, then for every $A = I_W + a \in \mathcal{R}_a(\mu, \nu)$, $a \in L^2_a(\mu, H)$, the innovation conjecture is wrong*

Proof: Let $A = I_W + a \in \mathcal{R}_a(\mu, \nu)$, $a \in L^2_a(\mu, H)$. First assume that V is $\nu-$invertible, so that $V \circ U = I_W$. Thus from the Lemma VI.4 we know that $A = U \circ Z^A$. Hence $\mathcal{F}^A \subset \mathcal{F}^{Z^A}$. Moreover, by definition $\mathcal{F}^{Z^A} \subset \mathcal{F}^A$, so that $\mathcal{F}^A = \mathcal{F}^{Z^A}$. Conversely, assume that $\mathcal{F}^A = \mathcal{F}^{Z^A}$. Then there is a measurable a mapping D adapted to $\mathcal{F}^\nu_\cdot$ (which is the filtration of the coordinate process) such that $A = D \circ Z^A$. Furthermore $Z^A \in \mathcal{R}(\mu, \mu)$ because of Corollary VI.5. Therefore

$$D\mu = A \circ Z^A \mu = A\mu << \mu$$

On the other hand:

$$\begin{aligned} 1 &= \mu\left(V \circ A = Z^A\right) \\ &= \mu\left(V \circ D \circ Z^A = Z^A\right) \\ &= Z^A\mu\left(V \circ D(w) = w\right) \\ &= \mu\left(V \circ D(w) = w\right) \end{aligned}$$

Hence D is an API which represents ν, and V is the almost sure right inverse of D . Thus the Proposition VI.11 yields the$\nu-$invertibility of V. By contraposition, we get the result. $\square$

VII

Stochastic invertibility on Wiener space for stochastic differential equations with dispersion

Abstract. **A general result. The Brownian transform. Invertibility of the Brownian transform. Invertible laws of diffusion. Quasi invariance with respect to an invertible law of stochastic differential equation. Entropy based criterion for laws absolutely continuous with respect to an invertible law of diffusion.**

1. Introduction

In the previous chapter, we extended the notion of invertibility to the Girsanov shifts of any probability which is absolutely continuous with respect to the Wiener measure, but not necessarily equivalent. The natural question which arises is whether this notion can still make sense when one considers a probability ν which is not absolutely continuous with respect to the Wiener measure μ. Since the notation $\frac{d\nu}{d\mu}$ is no more defined, we have to find another way to define it, instead of the property $\nu - a.s.$

$$\frac{d\nu}{d\mu} = \rho\left(-\delta^W v\right) \tag{1.65}$$

which we used to define the Girsanov shift $V = I_W + v$ in the previous chapters. As a matter of fact in the case of a probability $\nu << \mu$ on the classical Wiener space, we already noticed that it was equivalent to define $V := I_W + v$, where $v \in L^0_a(\nu, H)$ satisfies (1.65), or to define V to be such that (I_W, V) is a weak solution of

$$dX_t = dB_t - \dot{v}_t(X)\,dt; X_0 = 0 \tag{1.66}$$

on $(W, \mathcal{F}^\nu_\cdot, \nu)$. Hence, instead of considering a probability $\nu << \mu$, we can start from the equation (1.66), and define ν to be the law of a solution, and V (if it exists) to be such that (I_W, V) is a solution of (1.66) on $(W, \mathcal{F}^\nu, \nu)$: in that case, V is still the Girsanov shift associated with ν. Similarly for a probability ν which is the law of a solution to a stochastic differential equation of the shape

$$dX_t = \alpha_t(X)\,dB_t + \beta_t(X)\,dt; X_0 = x \tag{1.67}$$

we will assume that there is a unique V with the property that (I_W, V) is solution of (1.67) on $(W, \mathcal{F}^\nu_\cdot, \nu)$. In that case, we have $V : W \to W$ and $V\nu = \mu$ so that $V \in M_\nu\left((W, \mathcal{F}^\nu), (W, \mathcal{F}^\mu)\right)$. Hence we can still define a notion of $\nu-$invertibility (formally the same as in Chapter VI) for this V. However, this assumption does not seem sufficient to relate the invertibility of V to the existence of a unique strong solution. For that reason we have made some further non-degeneracy assumptions. However, even if it is necessary to relate the invertibility of V to the existence of a unique strong solution, this non-degeneracy assumption seems very strong, so that we have weaken it in the last section, and we have explained how our previous results naturally extend. Before going further we emphasize that if the law ν of the solution is such that $\nu << \mu$, the Brownian transform of ν is the associated Girsanov shift, but that in the general case, the Brownian transform is not a perturbation of the identity.

The structure of this Chapter is the following. In Section 2 we fix the main hypothesis and notations which enable us to handle with morphisms of probability spaces, and we show a result of representation. In Section 3 we define the Brownian transform of a law, under a further assumption. In Section 4, we define the invertibility of the Brownian transform, and we relate this result to the existence of a unique strong solution for the underlying stochastic differential equation. In Section 5 we study the structure of quasi-invariant flows with respect to a law whose Brownian transform is invertible. In Section 6 we make a further remark on the Malliavin calculus with respect to such laws. In Section 7 we investigate the local properties of the invertibility. We then extend the criterion based on the entropy to this framework (Section 8). In Section 9 we investigate some further extensions of the Brownian transform : its invertibility is then related to some pathwise uniqueness among solutions with a same law.

2. Preliminaries and notations

Let (W,H,μ) be the classical Wiener space $(C([0,1],\mathbb{R}^d)$ for a $d\in\mathbb{N})$, $t\to W_t$ the coordinate process, and $(\mathcal{F}_t)$ the associated filtration. We note $\mathcal{A}^{d,r}$ the set of the progressively measurable processes (with respect to the canonical filtrations) $(\alpha_s, s\in[0,1])$ with values in $\mathbb{R}^d\otimes\mathbb{R}^r$. Throughout these chapter we will assume that $\alpha\in\mathcal{A}_{d,d}$ and $\beta\in\mathcal{A}_{d,1}$, and we fix a $x\in\mathbb{R}^d$. Moreover we assume the set of the laws of the weak solutions to (1.67), which we note $\mathcal{P}(\alpha,\beta)$, to be not empty, and we call $(H0)$ this hypothesis. Let $\nu\in\mathcal{P}(\alpha,\beta)\neq\emptyset$, and π^W be the projection on the first coordinate of $W\times W$. Then, (see 169 and 98 of **[22]**) there is a $\widehat{V}\in L^0(\nu\otimes\mu,W)$ such that $\left(\pi,\widehat{V}\right)$ is a solution of (1.67) on $(W\times W,\mathcal{F}_t\times\mathcal{F}_t,\nu\otimes\mu)$. Hence we have $\nu\otimes\mu-a.s.$

$$W_{\cdot}\left(\pi^W\omega\right)=x+\int_0^{\cdot}\alpha_s\left(\pi^W\omega\right)d\widehat{V}_s(\omega)+\int_0^{\cdot}\beta_s\left(\pi^W\omega\right)ds$$

and $\widehat{V}(\nu\otimes\mu)=\mu$. We call V the canonical Brownian associated with ν. We note μ the Wiener measure W the Wiener space, and $(\mathcal{F}^\nu_t)$ the augmentation of the filtration generated by the coordinate process $t\to W_t$ with respect to a borelian probability ν on W, and we note $\mathcal{F}^\nu$ the completion of the borelian sigma field $\mathcal{B}(W)$ with respect to ν. We henceforth assume that $\nu\in\mathcal{P}(\alpha,\beta)$ and we note $\widehat{V}$ the associated canonical Brownian. Let $(\Omega,\mathcal{G},\mathbb{P})$ be a complete probability space we set $\widetilde{\Omega}:=\Omega\times W$, $\widetilde{\mathcal{G}}:=\mathcal{G}\otimes\mathcal{F}^\mu$, $\widetilde{\mathbb{P}}:=\mathbb{P}\otimes\mu$. We note $\pi:\widetilde{\Omega}\to\Omega$ (resp. $\widehat{\pi}:\widetilde{\Omega}\to W$) the projection on the first (resp. second) coordinate of $\widetilde{\mathcal{G}}$. From the very definition of the projection we have $\pi\in\mathcal{M}\left(\left(\widetilde{\Omega},\widetilde{\mathcal{G}}\right),(\Omega,\mathcal{G})\right)$ (see section 2 of Chapter VI for these notations) and $\widehat{\pi}\in\mathcal{M}\left(\left(\widetilde{\Omega},\widetilde{\mathcal{G}}\right),(W,\mathcal{F}^\mu)\right)$. We now consider a $U\in\mathcal{M}\left((\Omega,\mathcal{G}),(W,\mathcal{B}(W))\right)$ such that $U\mathbb{P}<<\nu$, so that we also have $U\in\mathcal{M}\left((\Omega,\mathcal{G}),(W,\mathcal{F}^\nu)\right)$, and $U\circ\pi\in\mathcal{M}\left((\Omega,\mathcal{G}),(W,\mathcal{F}^\nu)\right)$. We define the **lift** $\widetilde{U}$ **of** U on $\widetilde{\Omega}$ by

$$\boxed{\widetilde{U}:z\in\widetilde{\Omega}\to\left(U\circ\pi(z),\widehat{\pi}(z)\right)\in W\times W}$$

From the properties of the projections and since $U\mathbb{P}<<\nu$, we have $\widetilde{U}\widetilde{\mathbb{P}}<<\widetilde{\mu}$ and $\widetilde{U}\in\mathcal{M}\left(\left(\widetilde{\Omega},\widetilde{\mathcal{G}}\right),\left(\widetilde{W},\widetilde{\mathcal{F}^\nu}\right)\right)$. Since $\widehat{V}\widetilde{\nu}=\mu$ we have $\widehat{V}\in\mathcal{M}\left(\left(\widetilde{W},\widetilde{\mathcal{F}^\nu}\right),(W,\mathcal{F}^\mu)\right)$. It is straightforward to check that the equivalence class of $\widehat{V}\circ\widetilde{U}$ in $M_{\mathbb{P}}\left(\left(\widetilde{\Omega},\widetilde{\mathcal{G}}\right),(W,\mathcal{F}^\mu)\right)$ only depends on the equivalence class of of $\widehat{V}$ in $M_{\widetilde{\nu}}\left(\left(\widetilde{W},\widetilde{\mathcal{F}^\nu}\right),(W,\mathcal{F}^\mu)\right)$ and of $\widetilde{U}$ in $M_{\widetilde{\mathbb{P}}}\left(\left(\widetilde{\Omega},\widetilde{\mathcal{G}}\right),\left(\widetilde{W},\widetilde{\mathcal{F}^\nu}\right)\right)$. Hence, for a $U\in M_{\mathbb{P}}\left((\Omega,\mathcal{G}),(W,\mathcal{F}^\nu)\right)$ such that $U\mathbb{P}<<\nu$, $\widehat{V}\circ\widetilde{U}$ is well defined as an element of $M_{\widetilde{\mathbb{P}}}\left(\left(\widetilde{\Omega},\widetilde{\mathcal{G}}\right),(W,\mathcal{F}^\mu)\right)$. We assume henceforth that $\Omega=W$,$\mathcal{G}=\mathcal{F}^\nu$ for a $\nu\in\mathcal{P}(\alpha,\beta)$, and we still consider a measurable mapping $U\in M_\mu\left((W,\mathcal{F}^\mu),(W,\mathcal{F}^\nu)\right)$ such that $U\mu<<\nu$. In that case not only $\widehat{V}\circ\widetilde{U}$ is well defined but also $U\circ\widehat{V}$. Indeed let $U\in M_\mu\left((W,\mathcal{F}^\mu)(W,\mathcal{F}^\nu)\right)$ and let $\widehat{V}\in M_\nu\left(\left(\widetilde{W},\widetilde{\mathcal{F}^\nu}\right)(W,\mathcal{F}^\mu)\right)$ still denote the canonical Brownian associated with ν we defined in the introduction. It is easy to check that $U\circ\widehat{V}$ is also well defined as an element of $M_{\widetilde{\nu}}\left(\left(\widetilde{W},\widetilde{\mathcal{F}^\nu}\right)(W,\mathcal{F}^\nu)\right)$. To avoid any ambiguity, we note π^W (resp. $\widehat{\pi}^W$) instead of π (resp. of $\widehat{\pi}$) in the case $\Omega=W$. With these notations we have $\widetilde{\mathbb{P}}-a.s.$ $\pi^W\widetilde{X}=X\circ\pi$. We note $\Pi:(x,y)\in\Omega\times W\to(x,0)\in\Omega\times W$. In the case $\Omega=W$ we note Π^W. Since $\widehat{V}\circ\Pi^W$ only depends on the first coordinate, we can find a V such that $\widetilde{\nu}-a.s.$

$$\boxed{V\circ\pi=\widehat{V}\circ\Pi^W}$$

Note that if $\widehat{V}\circ\Pi^W=\widehat{V}$ there is clearly no more need to consider the embeddings on the product space, and we have $V\nu=\mu$ and $V\in M_\nu\left((W,\mathcal{F}^\nu),(W,\mathcal{F}^\mu)\right)$: in the next sections we shall focus on that case. Under $(H0)$, for a $\nu\in\mathcal{P}(\alpha,\beta)$, we first prove that for any U which represents ν, it is solution of (1.67) on an extended space. This result may be well known but it will be very usefull in the sequel, and it will motivate our non-degeneracy assumption :

Theorem VII.1. *Assume that $(H0)$ holds, and let $\nu\in\mathcal{P}(\alpha,\beta)$, and $\widehat{V}$ be the canonical Brownian associated with ν. Moreover, let $(\Omega,\mathcal{G},\mathbb{P})$ be a complete probability space, and let $U\in M_{\mathbb{P}}\left((W,\mathcal{G}),(W,\mathcal{F}^\nu)\right)$ be such that*

$$U\mathbb{P}=\nu$$

for a $\nu\in\mathcal{P}(\alpha,\beta)$. We note $\widetilde{U}\in M_{\widetilde{\mathbb{P}}}\left(\left(\widetilde{\Omega},\widetilde{\mathcal{G}}\right),\left(\widetilde{W},\widetilde{\mathcal{F}^\nu}\right)\right)$ the lift of U (see the notations). Further, let $\left(\widetilde{\mathcal{U}}_t\right)$ be the filtration generated by $t\to\widetilde{U}_t$, augmented with respect to $\widetilde{\mathbb{P}}$, and set

$$\widehat{R^U}:=\widehat{V}\circ\widetilde{U}$$

Then $\left(U\circ\pi,\widehat{R^U}\right)$ is a weak solution of (1.67) on $\left(\widetilde{\Omega},\widetilde{\mathcal{G}},\widetilde{\mathbb{P}}\right)$ with the filtration $\left(\widetilde{\mathcal{U}}_t\right)$. In particular $\widehat{R^U}$ is a $\left(\widetilde{\mathcal{U}}_t\right)$ Brownian motion on $\left(\widetilde{\Omega},\widetilde{\mathcal{G}},\widetilde{\mathbb{P}}\right)$

Proof: We first show that $\left(\widehat{R^U}_t\right)$ is a Brownian motion. Let $\theta_s\in C_b\left(\widetilde{W}\right)$ be $\widetilde{\mathcal{F}}^\nu_s$ measurable. Since $U\mathbb{P}=\nu$, we have $\widetilde{U}\widetilde{\mathbb{P}}=\widetilde{\nu}$. Hence for any $t>s$ we obtain

$$\begin{aligned}
E_{\widetilde{\mathbb{P}}}\left[\left(\widehat{R^U}_t-\widehat{R^U}_s\right)\theta_s\circ\widetilde{U}\right] &= E_{\widetilde{\mathbb{P}}}\left[\left(\widehat{V}_t\circ\widetilde{U}-\widehat{V}_s\circ\widetilde{U}\right)\theta_s\circ\widetilde{U}\right]\\
&= E_{\widetilde{U}\widetilde{\mathbb{P}}}\left[\left(\widehat{V}_t-\widehat{V}_s\right)\theta_s\right]\\
&= E_{\widetilde{\nu}}\left[\left(\widehat{V}_t-\widehat{V}_s\right)\theta_s\right]\\
&= 0
\end{aligned}$$

where the last equality holds since $\left(\widehat{V}_t\right)$ is a Brownian motion on $\left(\widetilde{W},\widetilde{\mathcal{F}^\nu},\widetilde{\nu}\right)$. Hence $\left(\widehat{R^U}_t\right)$ is a $\left(\widetilde{\mathcal{U}}_t\right)$martingale on $\left(\widetilde{\Omega},\widetilde{\mathcal{G}},\widetilde{\mathbb{P}}\right)$. From Paul Levy's criterion we just have to prove that $<\widehat{R^U}>_t=t$. For a $s\in[0,1]$ let $\theta_s\in C_b\left(\widetilde{W}\right)$ be $\widetilde{\mathcal{F}}^\nu_s$ measurable, and $t\geq s$. Since $<V>_t=t$ we have :

$$\begin{aligned}
E_{\widetilde{\mathbb{P}}}\left[\left(\left(\widehat{R^U}_t\right)^2-t\right)\theta_s\circ\widetilde{U}\right] &= E_{\widetilde{\mathbb{P}}}\left[\left(\left(\widehat{V}_t\circ\widetilde{U}\right)^2-t\right)\theta_s\circ\widetilde{U}\right]\\
&= E_{\widetilde{U}\widetilde{\mathbb{P}}}\left[\left(\left(\widehat{V}_t\right)^2-t\right)\theta_s\right]\\
&= E_{\widetilde{\nu}}\left[\left(\widehat{V}_s^2-s\right)\theta_s\right]\\
&= E_{\widetilde{U}\widetilde{\mathbb{P}}}\left[\left(\left(\widehat{V}_s\right)^2-s\right)\theta_s\right]\\
&= E_{\widetilde{\mathbb{P}}}\left[\left(\left(\widehat{V}_s\circ\widetilde{U}\right)^2-s\right)\theta_s\circ\widetilde{U}\right]\\
&= E_{\widetilde{\mathbb{P}}}\left[\left(\left(\widehat{R^U}_s\right)^2-s\right)\theta_s\circ\widetilde{U}\right]
\end{aligned}$$

which prove the first assertion. We still have to prove that $\left(U,\widehat{R^U}\right)$ is a solution on the announced space. As a matter of fact this is comes from the very definition of V, of π, and of $\widetilde{U}$ (which is such that $\pi^W\widetilde{U}=U\circ\pi$).

Indeed we have $\widetilde{\mathbb{P}} - a.s.$

$$\begin{aligned}
(U \circ \pi)_. &= W_. (U \circ \pi) \\
&= W_. \left(\pi^W \widetilde{U}\right) \\
&= x + \int_0^. \alpha_s \left(\pi^W \widetilde{U}\right) d\widehat{V}_s \circ \widetilde{U} + \int_0^. \beta_s \left(\pi^W \widetilde{U}\right) ds \\
&= x + \int_0^. \alpha_s (U \circ \pi) \, d\widehat{R^U}_s + \int_0^. \beta_s (U \circ \pi) \, ds
\end{aligned}$$

□

We will also need the following notations to handle with morphisms of probability space. If $(\Omega, \mathcal{G}, \mathbb{P})$ is a probability space, we note $\mathcal{G}^{\mathbb{P}}$ the completion of $\mathcal{G}$ with respect to $\mathbb{P}$. We recall that $\mathbb{P}$ extends uniquely on this sigma field, and we will note $\Theta^{\mathbb{P}}$ the negligible sets associated with $\mathbb{P}$. For two probabilities $\mathbb{P}$ (resp. $\mathbb{Q}$) defined on a space $(\Omega, \mathcal{G})$ (resp. on a space $\left(\widetilde{\Omega}, \widetilde{\mathcal{G}}\right)$) we set

$$\mathcal{R}^0 (\mathbb{P}, \mathbb{Q}) := \left\{ X \in \mathcal{M}\left((\Omega, \mathcal{G}), \left(\widetilde{\Omega}, \widetilde{\mathcal{G}}\right)\right) \middle| X\mathbb{P} = \mathbb{Q} \right\}$$

and

$$\mathcal{R} (\mathbb{P}, \mathbb{Q}) := \left\{ X \in M_{\mathbb{P}}\left(\left(\Omega, \mathcal{G}^{\mathbb{P}}\right), \left(\widetilde{\Omega}, \widetilde{\mathcal{G}}^{\mathbb{Q}}\right)\right) \middle| X\mathbb{P} = \mathbb{Q} \right\}$$

3. The Brownian transform

Assume that $(H0)$ holds, and let $\nu \in \mathcal{P}(\alpha, \beta)$. We will further assume that $\widehat{V} = V \circ \pi^W$, and that this has further properties (which may be seen as a non-degeneracy assumption of the dispersion), so that (1.67) has similar properties to the equations investigated the previous chapters. Roughly speaking this assumptions means that we will focus on stochastic differential equations such that the observed signal (i.e. the solution of the equation) carries much information than the associated noise, and that the way the information is incorporated does not depend on the underlying solution : at least physically, this hypothesis seams very reasonable. However, in the last section, we shall relax this hypothesis.

Definition VII.1. *Assume that $(H0)$ holds and let $\nu \in \mathcal{P}(\alpha, \beta)$, we call $(H1)$ the following assumption : there exists a $V \in \mathcal{M}((W, \mathcal{B}(W)), (W, \mathcal{B}(W)))$ such that for any probability space $(\Omega, \mathcal{G}, \mathbb{P})$ and any $X, B \in \mathcal{M}((\Omega, \mathcal{G}), (W, \mathcal{B}(W)))$ the following assertions are equivalent*

(i) (X, B) is a solution of (1.67) on $(\Omega, \mathcal{G}, \mathbb{P})$

(ii) $B \in \mathcal{R}^0(\mathbb{P}, \mu)$ and $\mathbb{P} - a.s.$

$$V \circ X = B \tag{3.68}$$

Whenever $(H1)$ holds, we define the Brownian transform of ν to be the equivalence class of V in $\mathcal{R}(\nu|\mu)$.

We now check briefly that this definition makes sense. First note that $V \in \mathcal{R}^0(\nu, \mu)$. Indeed by definition of ν there is a (X, B) which is solution of (1.67) on some probability space $(\Omega, \mathcal{G}, \mathbb{P})$ so that $X\mathbb{P} = \nu$ and of course $B\mathbb{P} = \mu$. Then (3.68) implies $V\nu = \mu$ which exactly means that $V \in \mathcal{R}^0(\nu, \mu)$. On the other hand, it obviously implies that $V \in \mathcal{M}((W, \mathcal{F}^\mu), (W, \mathcal{F}^\nu))$ so that it makes sense to consider its equivalence class in $\mathcal{R}(\nu, \mu)$. We now check the uniqueness, and we investigate the main properties of V. Note that, with the notations of Theorem VII.1, the lift of I_W is $\widetilde{I_W} = \pi^W \times \widehat{\pi^W}$, so that $\widehat{V} \circ \widetilde{I_W} = \widehat{V}$ and $I_W \circ \pi^W = \pi^W$. Hence since ν is the law of a solution and $I_W \nu = \nu$, Theorem VII.1 implies that $(\pi^W, \widehat{V})$ is solution of (1.67) on $(W \times W, \nu \otimes \mu)$. Together with (3.68) it yields yields $\nu \otimes \mu$ a.s.

$$V \circ \pi^W = \widehat{V}$$

so that $\nu \otimes \mu$ a.s.

$$\pi^W = x + \int_0^\cdot \alpha_s \circ \pi^W dV_s \circ \pi^W + \int_0^\cdot \beta_s \circ \pi^W ds$$

where $V \circ \pi^W$ is a Brownian motion with respect to the filtration generated by π^W. This is equivalent to following main property of V : (I_W, V) is a solution of (1.67) on $(W, \mathcal{F}^\nu, \nu)$ with respect to the filtration generated by the coordinate process $(\mathcal{F}^\nu_\cdot)$. In particular we have the following which will be useful to achieve the calculus $\nu - a.s.$

$$I_W = x + \int_0^\cdot \alpha_s dV_s + \int_0^\cdot \beta_s ds$$

Note that the ν almost sure uniqueness is a straightforward consequence of (3.68). Indeed any $\widetilde{V}$ with the same property is such that $(I_W, \widetilde{V})$ is a solution of (1.67) on $(W, \mathcal{F}^\nu, \nu)$ so that (3.68) yields $\nu - a.s.$ $\widetilde{V} = V$. Hence the Brownian transform is well defined, and unique ν almost surely.

Until the end of this section, we check some technical details which are thinner than what we will really need. Indeed in the case of invertibility, which we will investigate latter, all these facts become trivial. Specifically, for the sake of completeness, we check that some inverse images of the Brownian transform are generally well defined.

Proposition VII.1. *Let $V \in \mathcal{R}^0(\nu, \mu)$, set $\mathcal{F}^V := V^{-1}(\mathcal{F}^\mu)$ and further assume that there is a $(\Omega, \mathcal{G}, \mathbb{P})$, a $X \in \mathcal{M}((\Omega, \mathcal{G}), (W, \mathcal{B}(W)))$ and a $B \in \mathcal{R}^0(\mathbb{P}, \mu)$ such that $\mathbb{P} - a.s.$*

$$V \circ X = B \tag{3.69}$$

Then we have

$$X \in \mathcal{M}\left(\left(\Omega, \mathcal{G}^{\mathbb{P}}\right), \left(W, \mathcal{F}^V\right)\right)$$

Proof: Since $\mathbb{P} - a.s.$, $V \circ X = B$ and $B^{-1}(\mathcal{B}(W)) \subset \mathcal{G}$, we obtain

$$\begin{aligned}
X^{-1}(\mathcal{F}^V) &= X^{-1}\sigma\left(V^{-1}\left(\mathcal{B}(W) \bigcup \Theta^\mu\right)\right) \\
&= \sigma\left(X^{-1}\left(V^{-1}\left(\mathcal{B}(W) \bigcup \Theta^\mu\right)\right)\right) \\
&= \sigma\left((V \circ X)^{-1}\left(\mathcal{B}(W) \bigcup \Theta^\mu\right)\right) \\
&= \sigma\left(B^{-1}\left(\mathcal{B}(W) \bigcup \Theta^\mu\right)\right) \\
&= \sigma\left(B^{-1}(\mathcal{B}(W)) \bigcup \left(B^{-1}(\Theta^\mu)\right)\right) \\
&\subset \sigma\left(\mathcal{G} \cup \left(B^{-1}(\Theta^\mu)\right)\right)
\end{aligned}$$

On the other hand, since $B\mathbb{P} = \mu$, $B^{-1}(\Theta^\mu) \subset \Theta^{\mathbb{P}}$. Hence

$$X^{-1}(\mathcal{F}^V) \subset \sigma\left(\mathcal{G} \cup \Theta^{\mathbb{P}}\right) = \mathcal{G}^{\mathbb{P}}$$

□

The following fact is clear, however it seams necessary to stress it explicitly for the reader's convenience

Proposition VII.2. *For any $V \in \mathcal{R}(\nu, \mu)$, V can be seen either as an element of*

$$M_\nu((W, \mathcal{F}^\nu), (W, \mathcal{F}^\mu))$$

or as an element of the subspace $M_{\nu|_{\mathcal{F}^V}}\left((W, \mathcal{F}^V), (W, \mathcal{F}^\mu)\right)$ where $\mathcal{F}^V = V^{-1}(\mathcal{F}^\mu)$, and where $\nu|_{\mathcal{F}^V}$ is the trace of ν on $\mathcal{F}^V$. Moreover $V \in \mathcal{R}(\nu|_{\mathcal{F}^V}, \mu)$

We recall that since $I_W \in \mathcal{R}\left(\nu, \nu|_{\mathcal{F}^V}\right)$, it defines an injection

$$M_{\nu|_{\mathcal{F}^V}}\left(\left(W, \mathcal{F}^V\right), \left(W, \mathcal{F}^\mu\right)\right) \to M_\nu\left(\left(W, \mathcal{F}^\nu\right), \left(W, \mathcal{F}^\mu\right)\right)$$

Hence we will sometimes identify V with an element of this subspace. The next proposition is thinner than what we will really need. We now investigate that the following pull-backs are generally well defined :

Proposition VII.3. *Assume that $(H0)$ holds, and let $\nu \in \mathcal{P}(\alpha, \beta)$ be such that $(H1)$ holds. Further note V the associated Girsanov transform, $\mathcal{F}^V = V^{-1}\left(\mathcal{F}^\mu\right)$, and let $U \in M_\mu\left(\left(W, \mathcal{F}^\mu\right), (Y, \mathcal{B})\right)$, where $(Y, \mathcal{B})$ is a measurable space. For any (X, B) which is solution of (1.67) on a probability space $(\Omega, \mathcal{G}, \mathbb{P})$, we then have that X is well defined as an element of $\mathcal{R}\left(\mathbb{P}, \nu|_{\mathcal{F}^V}\right)$. As a consequence we also have the following :*

(i) $V \circ X$ (resp. $U \circ B$) is well defined as an element of $M_\mathbb{P}\left(\left(\Omega, \mathcal{G}^\mathbb{P}\right), \left(W, \mathcal{F}^\mu\right)\right)$ (resp. as an element of $M_\mathbb{P}\left(\left(\Omega, \mathcal{G}^\mathbb{P}\right), (Y, \mathcal{B})\right)$) i.e. its is measurable with respect to the related sigma fields and its equivalence class only depends of the equivalence class of $X \in M_\mathbb{P}\left(\left(\Omega, \mathcal{G}^\mathbb{P}\right), \left(W, \mathcal{F}^V\right)\right)$ (resp. of B in $M_\mathbb{P}\left(\left(\Omega, \mathcal{G}^\mathbb{P}\right), \left(W, \mathcal{F}^\mu\right)\right)$) and of the equivalence class of $V \in M_{\nu|_{\mathcal{F}^V}}\left(\left(W, \mathcal{F}^V\right), \left(W, \mathcal{F}^\mu\right)\right)$ (resp. of $U \in M_\mu\left(\left(W, \mathcal{F}^\mu\right), (Y, \mathcal{B})\right)$).

(ii) $U \circ (V \circ X)$ is well defined as an element of $M_\mathbb{P}\left(\left(\Omega, \mathcal{G}^\mathbb{P}\right), (Y, \mathcal{B})\right)$ i.e. its equivalence class only depends of the equivalence class of $U \in M_\mu\left(\left(W, \mathcal{F}^\mu\right), (Y, \mathcal{B})\right)$, of $V \in M_{\nu|_{\mathcal{F}^V}}\left(\left(W, \mathcal{F}^V\right), \left(W, \mathcal{F}^\mu\right)\right)$ (and so in $M_\nu\left(\left(W, \mathcal{F}^\nu\right), \left(W, \mathcal{F}^\mu\right)\right)$), and of $X \in M_\mathbb{P}\left(\left(\Omega, \mathcal{G}^\mathbb{P}\right), \left(W, \mathcal{F}^V\right)\right)$.

Proof: Note that (i) clearly yields (ii). On the other hand to prove (i), it is sufficient to check that $X \in \mathcal{R}\left(\mathbb{P}, \nu|_{\mathcal{F}^V}\right)$ (resp. that $B \in \mathcal{R}(\mathbb{P}, \mu)$). Note that $B \in \mathcal{R}(\mathbb{P}, \mu)$ is obvious, and that we already proved in Proposition VII.2 that $X \in M_\mathbb{P}\left(\left(\Omega, \mathcal{G}^\mathbb{P}\right), \left(W, \mathcal{F}^V\right)\right)$. In this proposition we see X as an element of this space, and we just have to prove that $X\mathbb{P} = \nu|_{\mathcal{F}^V}$. Let $A \in \mathcal{F}^V$, and let $D \in \mathcal{F}^\mu$ be such that $A = V^{-1}(D)$. Then, since $\mathbb{P} - a.s.$ $V \circ X = B$ and $B\mathbb{P} = \mu$, we obtain

$$\begin{aligned} X\mathbb{P}(A) &= \mathbb{P}\left(X^{-1}(A)\right) \\ &= \mathbb{P}\left(X^{-1}\left(V^{-1}(D)\right)\right) \\ &= \mathbb{P}\left((V \circ X)^{-1}(D)\right) \\ &= \mathbb{P}\left(B^{-1}(D)\right) \\ &= B\mathbb{P}(D) \\ &= \mu(D) \end{aligned}$$

On the other hand we have $V\nu = \mu$ so that :

$$\begin{aligned} X\mathbb{P}(A) &= \mu(D) \\ &= V\nu(D) \\ &= \nu\left(V^{-1}(D)\right) \\ &= \nu(A) \end{aligned}$$

□

4. Invertibility of the Brownian transform

Definition VII.2. *Assume that $(H0)$ holds, let $\nu \in \mathcal{P}(\alpha, \beta)$ be such that $(H1)$ holds, and let V be the Brownian transform of ν. We say that V is $\nu-$right invertible if there is a $U \in M_\mu\left(\left(W, \mathcal{F}^\mu\right), \left(W, \mathcal{F}^\nu\right)\right)$ which is adapted to $\mathcal{F}^\mu_\cdot$, such that $U\mu << \nu$ and $\mu - a.s.$*

$$V \circ U = I_W$$

Moreover V is said to be $\nu-$invertible if and only if it is right invertible and $\nu - a.s.$

$$U \circ V = I_W$$

For short we will say that $\nu \in \mathcal{P}(\alpha,\beta)$ is invertible if and only if $(H1)$ holds for ν, and if the Brownian transform V of ν is invertible

Note that with the same proof as in the Brownian case (See Chapter VI), it is easy to check that the left-invertibility (i.e. $\nu-$a.s. $U \circ V = I_W$) implies the right invertibility, and that we still have

Proposition VII.4. *Assume that $(H0)$ holds, let $\nu \in \mathcal{P}(\alpha,\beta)$ be such that $(H1)$ holds, and let V be the Brownian transform of ν. Further assume that V is right invertible with inverse U. Then the following assertions are equivalent*

(i) V is $\nu-$invertible with inverse U
(ii) $U\mu = \nu$

Many good properties of invertible diffusions follows from the following proposition, which comes with the same proof as Proposition VI.10 and Proposition VI.11

Proposition VII.5. *Consider the same hypothesis as in Definition VII.2, and take the same notations. Further assume that V is invertible with inverse U, then we have*

$$\mathcal{F}^{\mu}_{\cdot} = \mathcal{F}^{U}_{\cdot}$$

and

$$\mathcal{F}^{\nu}_{\cdot} = \mathcal{F}^{V}_{\cdot}$$

Remark VII.1. *Let V be the Brownian transform of a probability ν under the two above assumptions. It can be seen easily that the right invertibility is sufficient to get a left invertibility on $(W, \mathcal{F}^V, \nu|_{\mathcal{F}^V})$. Specifically, let $\widetilde{V} \in \mathcal{M}((W,\mathcal{B}(W),(W,\mathcal{B}(W)))$ be in the equivalence class of V and let $\widetilde{U} \in \mathcal{M}((W,\mathcal{B}(W)),(W,\mathcal{B}(W))$ be such that $\mu - a.s.$*

$$\widetilde{V} \circ \widetilde{U} = I_W$$

where I_W is seen as an elements of $\mathcal{M}((W,\mathcal{B}(W)),(W,\mathcal{B}(W))$. For instance such a $\widetilde{U}$ can be sometimes obtained from fixed point techniques, or variational techniques. By Proposition VII.2 and Proposition VII.3, if we set U to be the equivalence class of $\widetilde{U}$ in $\mathcal{R}(\mu, \nu|_{\mathcal{F}^V})$, and if we see V as an element of $\mathcal{R}(\nu|_{\mathcal{F}^V}, \mu)$, the same proof implies that $\nu|_{\mathcal{F}^V} - a.s.$

$$U \circ V = I_W$$

where I_W is now seen as an element of $\mathcal{R}(\nu|_{\mathcal{F}^V}, \nu|_{\mathcal{F}^V})$ and $\mu - a.s.$

$$V \circ U = I_W$$

where I_W is seen as an element of $\mathcal{R}(\mu,\mu)$. Of course, the left invertibility in this relaxed sense also implies the right invertibility in this relaxed sense. With other words the existence of a strong solution U on the Wiener space (i.e. $\mu - a.s.$ $V \circ U = I_W$, and (U_t) adapted to the canonical filtration) is sufficient to get an isomorphism of filtered probability space between $(W, \mathcal{F}^V_{\cdot}, \nu|_V)$ (where $\mathcal{F}^V_{\cdot} = V^{-1}(\mathcal{F}^{\mu}_{\cdot})$) and the Wiener space : this can be used to generalize the results of this whole chapter to the spaces $(W, \mathcal{F}^V_{\cdot}, \nu|_V)$ where the associated equation has a strong solution (not necessarily unique) on the Wiener space. Finally note that this relaxed invertibility (which won't be considered below) is equivalent with Definition VII.2 if and only if (I_W, V) is a strong solution on $(W, \mathcal{F}^{\nu}_{\cdot}, \nu)$.

Lemma VII.1. *Assume that* $(H0)$ *holds, let* $\nu \in \mathcal{P}(\alpha, \beta)$ *be such that* $(H1)$ *holds, and let* V *be the Brownian transform of* ν. *Further assume that* V *is* $\nu-$*invertibile with inverse* U. *Then for nay* (X, B) *which is solution of (1.67) on a probability space* $(\Omega, \mathcal{G}, \mathbb{P})$, X *is well defined as an element of* $\mathcal{R}(\mathbb{P}, \nu)$, *and we have* $\mathbb{P} - a.s.$

$$(U \circ V) \circ X = U \circ (V \circ X)$$

and

$$(V \circ U) \circ B = V \circ (U \circ B)$$

Proof: First note that since U is invertible $U \in \mathcal{R}(\mu, \nu)$. Hence, together with Proposition VII.5 (which implies $\mathcal{F}^V = \mathcal{F}^\nu$), Proposition VII.3 implies that X is well defined as an element of $M_{\mathbb{P}}\left(\left(\Omega, \mathcal{G}^{\mathbb{P}}\right), (W, \mathcal{F}^\nu)\right)$. Furthermore from Proposition VII.3 we also know that $\mathcal{F}^V = \mathcal{F}^\nu$ implies $X\mathbb{P} = \nu$, so that we have $X \in \mathcal{R}(\mathbb{P}, \nu)$. On the other hand we know that $U \circ V$ is well defined as an element of $\mathcal{R}(\nu, \nu)$ and that its equivalence class only depends on the equivalence class of U in $\mathcal{R}(\mu, \nu)$ and of the equivalence class of V in $\mathcal{R}(\nu, \mu)$. Hence $(U \circ V) \circ X$ is well defined as an element of $\mathcal{R}(\mathbb{P}, \nu)$ and its equivalence class only depends on the equivalence class of U, V and X in the spaces we have recalled in this paragraph. On the other hand, Proposition VII.3 yields the similar result for $U \circ (V \circ X)$, with the same associated spaces since $\mathcal{F}^V = \mathcal{F}^\nu$. This is proves the first assertion. Similarly the fact that $(V \circ U) \circ B = V \circ (U \circ B)$, is directly implied by $B \in \mathcal{R}(\mathbb{P}, \mu)$, $U \in \mathcal{R}(\mu, \nu)$, and on the other hand by $U \circ B \in \mathcal{R}(\mathbb{P}, \nu)$. □

Theorem VII.2. *Assume that* $(H0)$ *holds, and let* $\nu \in \mathcal{P}(\alpha, \beta)$*be such that* $(H1)$ *holds. Then the following assertions are equivalent:*

(i) The equation (1.67) has a unique strong solution
(ii) The Brownian transform of ν *is* $\nu-$*invertible.*

In that case, we will sometimes say shortly that ν, *which is then the law of the unique strong solution, is invertible.*

Proof: Assume that (1.67) has a unique strong solution, and note U the strong solution on $(W, \mathcal{F}^\mu, \mu)$ with the canonical Brownian $t \to W_t$. Then $(H1)$ yields

$$V \circ U = I_W$$

$\mu - a.s.$ On the other hand, the uniqueness of the law implies $U\mu = \nu$, so that U is the ν inverse of V. Conversely , assume that V is $\nu-$invertible with inverse V. Let B be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. From Lemma VII.1 we obtain $\mathbb{P} - a.s.$

$$B = (V \circ U) \circ B = V \circ (U \circ B)$$

so that $(H1)$ implies that $(U(B), B)$ is solution of (1.67). Furthermore, let (X, B) be a solution of (1.67) on $(\Omega, \mathcal{F}, \mathbb{P})$. Then $(H1)$ yields $\mathbb{P} - a.s.$

$$V \circ X = B$$

Together with Lemma VII.1, it yields

$$X = (U \circ V) \circ X = U \circ (V \circ X) = U(B)$$

which is the result.

5. The structure of quasi-invariant flows

Assume that $(H0)$ holds and that $(H1)$ holds for a $\nu \in \mathcal{P}(\alpha,\beta)$. Let V be the associated Brownian transform : in this whole section and in the next one we assume that V is $\nu-$invertible with an inverse U. Then V may be seen as an isomorphism of filtered probability space and to any transformation $\tau : W \to W$ of the Wiener space under which $\tau\mu << \mu$ we can associate the transformation $\xi : W \to W$ by $\xi := U \circ \tau \circ V$ which is well defined. In that case we have $\xi\nu << \nu$. Conversely for a $\xi : W \to W$ such that $\xi\nu << \nu$ we can set $\tau = V \circ \xi \circ U$, and we have $\tau\mu << \mu$. Moreover Proposition VII.5 implies that τ is $\mathcal{F}^{\mu}_{\cdot}$ adapted if and only if ξ is $\mathcal{F}^{\nu}_{\cdot}$ adapted. That is, V generates a $1:1$ correspondence between the (adapted) quasi-invariant flows on $(W,\mathcal{F}^{\nu},\nu)$ and $(W,\mathcal{F}^{\mu},\mu)$. We first use this correspondence to study the structure of quasi-invariant flows on $(W,\mathcal{F}^{\nu},\nu)$. In the next section, we explain how to build naturally a weak Sobolev derivative with respect to invertible laws. Note that the idea to introduce an analogous of the Girsanov drift to study the absolute continuity of probabilities absolutely continuous with respect to the law of a diffusion, as well as the idea to extend Föllmer's formulation of the relative entropy to this context is not new (for instance ([**5**]) and the references therein). On the other hand, since we learned the relation between the main result of [**46**] and of [**3**], we noticed that some results of this this section has to be compared with the Section 4 of [**3**]. Moreover, the study of quasi invariant flows for laws absolutely continuous with respect to the law of a unique strong solution must be well known. However, our approach which focusses on isomorphisms of probability spaces through the Brownian transform seems to be both concise and enlightening, and they extend to the weaker hypothesis we will set in the last section.

Proposition VII.6. *Let ν be the law of a stochastic differential equation which satisfies $(H0)$ and $(H1)$. Further assume that the associated Brownian transform V^{ν} is $\nu-$invertible. Then for any probability $\eta << \nu$, there is a unique $v \in L^0_a(\eta,H)$ such that $\eta - a.s.$*

$$\frac{d\eta}{d\nu} = \rho\left(-\delta^{V^{\nu}} v\right) \tag{5.70}$$

and

$$V^{\eta} = V^{\nu} + v$$

is a Brownian motion on $(W,\mathcal{F}^{\eta}_{\cdot},\eta)$. We say that v is the Girsanov drift associated with η with respect to ν

Proof: Since V^{ν} is invertible $\mathcal{F}^{V}_{\cdot} = \mathcal{F}^{\nu}$. Hence the martingale representation theorem shows the existence of a $\alpha \in L^0_a(\nu,H)$ such that $\nu - a.s.$

$$\frac{d\eta}{d\nu} = 1 + \int_0^1 \dot{\alpha}_s dV^{\nu}_s$$

Let $L_s = E_{\nu}\left[\frac{d\eta}{d\nu}\middle|\mathcal{F}^{\nu}\right]$, and set $\dot{v}_s := -\frac{\dot{\alpha}_s}{L_s}$. The same argument as in the Brownian case shows that $v \in L^0_a(\eta,H)$. From the Girsanov theorem $\delta^{V^{\nu}} v$, which is seen as an element of $L^0_a(\eta)$, is well defined as the stochastic integral on $(W,\mathcal{F}^{\eta},\eta)$ of v with respect to V^{ν} which is a semi martingale under this probability. Hence we get (5.70). From the Girsanov theorem, we directly obtain the last part of the claim. □

Proposition VII.7. *Let ν be the law of a stochastic differential equation which satisfies $(H0)$ and $(H1)$. Further assume that the associated Brownian transform V^{ν} is $\nu-$invertible. For any $\eta << \nu$, there is a $v \in L^0_a(\eta,H)$ such that η is the law of the stochastic differential equation*

$$dX_t = \alpha_t(X)\,dB_t + (\beta_t(X) - \alpha_t(X)[\dot{v}_t \circ X])\,dt; X_0 = x \tag{5.71}$$

Moreover, the equation (5.72) has a Brownian transform $V^{\eta} := V^{\nu} + v$. In particular, for any $U : \Omega \to W$, $\mathcal{F}^{\mathbb{P}}/\mathcal{B}(W)$ measurable, such that $U\mathbb{P} << \nu$, where $(\Omega,\mathcal{F}^{\mathbb{P}},\mathbb{P})$ is a complete probability space, we have that $(U, V^{\eta} \circ U)$ is solution of :

$$dX_t = \alpha_t(X)\,dB_t + (\beta_t(X) - \alpha_t(X)[\dot{v}_t \circ X])\,dt; X_0 = x \tag{5.72}$$

on $(\Omega, \mathcal{F}^{\mathbb{P}}, \mathbb{P})$ *with respect to the filtration generated by* U*, where* v *is the Girsanov drift associated with* $U\mathbb{P}$ *with respect to* ν*.*

Proof: From the very definition of V^ν, and since $\eta << \nu$, we have $\eta - a.s$

$$I_W = x + \int_0^{\cdot} \alpha_t dV_t^\nu + \int_0^{\cdot} \beta_t dt \tag{5.73}$$

Let v be the drift defined in Proposition VII.6 and $V^\eta := V^\nu + v$. Then, the last equation directly implies (5.73), which is the first part of the claim. Theorem VII.1 achieves the proof. $\square$

Lemma VII.2. *Let* ν *be an invertible law of (1.67) whose Brownian transform is noted* V*. Further assume that* V *is* $\nu-$*invertible, and let* U *be its inverse. Then, for any* $a \in L_a^0(\nu, H)$ *and for any* $(\mathcal{F}_t^\nu)$ *Brownian motion* B *on* $(W, \mathcal{F}^\nu, \nu)$*, and set* $A := U \circ (B + a)$*. Then* (A, B) *is solution of*

$$dA_t = \alpha_t(A)\, dB_t + (\beta_t(A) + \alpha_t(A)[\dot{a}_t])\, dt; A_0 = x \tag{5.74}$$

on $(W, \mathcal{F}_{\cdot}^\nu, \nu)$*. Moreover* $A\nu << \nu$*, and* $A\nu \sim \nu$ *if and only if* $E_\nu\left[\rho\left(-\delta^B a\right)\right] = 1$*.*

Proof: Let $A := U(B + a)$, where U is the $\nu-$inverse of V. Then $A = U \circ \tau_a \circ V$ where $\tau_a = \widetilde{B} + \widetilde{a}$ and $\widetilde{B} = B \circ U$ and $\widetilde{a} = a \circ U$. The hypothesis on a and B together with the fact that U is the inverse of V imply that $\widetilde{a} \in L_a^0(\mu, H)$ and that $\widetilde{B}$ is a $(\mathcal{F}_t^\mu)$ Brownian motion on $(W, \mathcal{F}^\mu, \mu)$. Hence (see **[50]**) $\tau_a\mu << \mu$ i.e. $\tau_a \circ V\nu << \mu$ so that $A\nu := U \circ \tau_a \circ V\nu << U\mu = \nu$. Moreover $E_\nu\left[\rho\left(-\delta^B a\right)\right] = 1$ implies $E_\mu\left[\rho\left(-\delta^{\widetilde{B}}\widetilde{a}\right)\right] = 1$ which is well known to imply $\tau_a\mu \sim \mu$. We only have to check that (A, B) solves (5.74). However, since $A\nu << \nu$ the pullbacks are well defined and the result directly follows from the very definition of U. $\square$

6. Remark : weak derivative along the quasi-invariant flows

In this section, we still work under the hypothesis of the previous section and we show how to extend Malliavin calculus to invertible diffusion. Since it is very simple and natural within the framework of stochastic analysis, the following results must be well known and that's the reason why we just put it there as a remark for the sake of completeness, instead of to develop it completely. We now recall the following facts (See the Theorem 2.1 of **[8]** and see **[10]** where they use similar ideas for higher purposes than us, and in finer framework). Let $\phi : W \to W$ be such that $\phi\mu << \mu$, then from the result of the last chapter, we have $\phi = R^\phi + \widehat{\phi}$ where R^ϕ is a $\mathcal{F}_{\cdot}^\phi$ Brownian motion (i.e. with respect to the filtration generated by ϕ). If we further assume that ϕ is invertible in the sense that $\left(\mathcal{F}_t^\phi\right) = (\mathcal{F}_t^\mu)$ for any $t \in [0, 1]$, the martingale representation theorem together with Paul Levy's criterion imply that there is a $O(d)$ ($d \times d$ orthogonal matrix) valued adapted process $t \to O_t$ and a $h \in L_a^0(\mu, H)$ such that $\mu - a.s.$

$$\phi = \int_0^{\cdot} O_t dW_t + \int_0^{\cdot} \dot{h}_t dt$$

It is usual to consider following curves $\epsilon \to \phi^\epsilon[q, h]$ for a $s \to q_s$ which is adapted process with values in $so(d)$(antisymetric $d \times d$ matrix) vanishing at 0 and $h \in L_a^0(\mu, H)$: we set $\mu - a.s.$

$$\phi^\epsilon[q, h] = \int_0^{\cdot} \exp(\epsilon q_t)\, dW_t + \epsilon \int_0^{\cdot} \dot{h}_t dt$$

For a fixed h, q, the tangent process to this curve is then defined by $d\xi_t = q_t dW_t + \dot{h}_t dt$. Since the Brownian transform of ν is assumed to be invertible, the associated flows on $(W, \mathcal{F}^\nu, \nu)$ are given (see the proof of Lemma VII.2) as being solutions of

$$\psi^\epsilon[q, h] = x + \int_0^{\cdot} \alpha_t(\psi^\epsilon[q, h]) \exp(\epsilon q_t)\, dV_t + \int_0^{\cdot} \left(\epsilon\alpha(\psi^\epsilon[q, h])\left[\dot{h}_t\right] + \beta(\psi^\epsilon[q, h])\right) dt$$

where $h \in L_a^0(\nu, H)$ and $t \to q_t$ is $(\mathcal{F}_t^\nu)$ adapted with values in $so(d)$ vanishing at 0. We now explain how the Malliaivin derivative can be built on $(W, \mathcal{F}^\nu, \nu)$. If we set for a deterministic q, $R_q^\epsilon = \psi^\epsilon[q, 0]$ (resp. for a deterministic h $T_h^\epsilon = \psi^\epsilon[h, 0]$) we get a $1-$parameter sub-group of rotations (resp. a 1 parameter subgroup of translations) on $(W, \mathcal{F}^\nu, \nu)$. We now show how the Malliavin derivative can be transported on that space : for convenience of notations we build it for the translations, but this construction can be directly transported to rotations, or in the direction of a general adapted tangent process. Note that the polynomial of the form $F = f\left(\delta^V h_1, ..., \delta^V h_n\right)$ (where $h_1, ..., h_n \in H$ and f is a polynomial function) are dense in all the $L^p(\nu)$. We define the derivative on such functional in the direction $h \in H$ by :

$$\nabla_h^\nu F = \frac{d}{d\epsilon} F \circ T_h^\epsilon|_{\epsilon=0} = \left(\frac{d}{d\epsilon} f\left(\delta^W h_1, ..., \delta^W h_n\right) \circ \tau_{\epsilon h}|_{\epsilon=0}\right) \circ V$$

where $\tau_h = I_W + h$ and where we used that $T_h^\epsilon = U \circ \tau_{\epsilon h} \circ V$. i.e. $\nabla_h^\nu F = \sum_i h_i (\partial_i f)\left(\delta^V h_1, ..., \delta^V h_n\right) = (\nabla_h F \circ U) \circ V$ where ∇ denotes the weak Sobolev derivative on the Wiener space $(W, \mathcal{F}^\mu, \mu)$ which was defined at the beginning of the previous chapter. It is straightforward to check that ∇^ν is closable with dense domain and we note δ^ν the associated adjoint (we have $\delta^\nu F = (\delta(F \circ U)) \circ V$). Moreover the associated Sobolev space are easily seen to be the image of the Sololev space associated with ∇ on the Wiener space $\mathbb{D}_{p,q}(\nu, E) := V(\mathbb{D}_{p,k}(\mu, E))$. Since $\mathcal{F}_t^V = \mathcal{F}_t^\nu$ for any t, it is also straightforward to check that the Clark-Ocone formula also holds with respect to ν. Specifically we have

$$F = E_\nu[F] + \delta^\nu \pi^\nu \nabla^\nu F$$

for any $F \in \mathbb{D}_{2,1}(\nu)$ where $\pi^\nu u = \int_0^\cdot E_\nu[\dot{u}_s | \mathcal{F}_s^\nu] ds$. Concerning the generalized Wiener functional (see **[45]** or **[22]**,**[56]**,**[41]** and **[52]**,**[57]** for nice applications) we can still define some distributions in the sense of Watanabe with respect to ν. The space of the test function can be defined by

$$\mathbb{D}(\nu) = \cap_{p>1, k>0} \mathbb{D}_{p,k}(\nu)$$

we can still endow with the natural Fréchêt structure, and we can define the distributions of $(W, \mathcal{F}^\nu, \nu)$ to be the elements of the dual space noted $\mathbb{D}'(\nu)$. The divergence in the sense of the distribution $D^{*\nu}$ (resp. the derivative in the sense of distributions D^ν) on $(W, \mathcal{F}^\nu, \nu)$ can still be defined as being the transpose of ∇^ν (resp. the transpose of δ^ν) : of course $D^{*\nu}$ extends δ^ν and D^ν extends ∇^ν.

7. Local properties

We still use the fact that invertible laws ν induces an isomorphism of filtered probability space between $(W, \mathcal{F}^\nu, \nu)$ to transport the results on the Wiener space.

Proposition VII.8. *Let ν be an invertible law of (1.67) whose Brownian transform is noted V^ν, and let η be a probability absolutely continuous with respect to ν. We note V^η the Brownian transform of η and v the Girsanov drift associated with η with respect to ν which is given in Proposition VII.6. Further assume that there is a sequence (τ_n) of optional times on $(W, \mathcal{F}^\eta, \eta)$ such that $\eta - a.s.$ $\tau_n \uparrow 1$ with the further property that for any $n \in \mathbb{N}$ the law η^n defined by*

$$\eta^n := exp\left(-\delta^{V^\eta}((\pi_{\tau_n} - I_H) v) - \frac{|(\pi_{\tau_n} - I_H) v|_H^2}{2}\right).\eta$$

is invertible. Then η is invertible.

Proof: We note U^ν the $\nu-$inverse of the Brownian transform V^ν of ν, and for any $n \in \mathbb{N}$ we set $V^n = V^\nu + \pi_{\tau_n} v$ so that V^n is the Brownian transform of η^n. We then have $\mu - a.s.$

$$V^n \circ U^\nu = I_W + \pi_{\tau_n \circ U^\nu}(v \circ U^\nu)$$

and

$$V^\eta \circ U^\nu = (I_W + v \circ U^\nu) \circ V^\nu$$

Moreover by taking the sequence $(\tau_n \circ U^\nu)$, the hypothesis directly imply that the shift $I_W + v \circ U^\nu$ is locally invertible in the sense of Chapter VI, and hence invertible. Let $\widetilde{U}$ denote its inverse. On the other hand $V^\eta = (I_W + v \circ U^\nu) \circ V^\nu$. Therefore it is $\eta-$invertible with inverse $U^\eta := U^\nu \circ \widetilde{U}$. □

8. Criterion of invertibility

Proposition VII.9. *Let ν be the law of a stochastic differential equation which satisfies $(H0)$ and $(H1)$. Further assume that the associated Brownian transform V^ν is $\nu-$invertible. Then, for any $\eta << \nu$*

$$2H(\eta|\nu) = E_\eta\left[|v|_H^2\right]$$

where v is the Girsanov drift associated with η with respect to ν.

Proof: By the same localization argument as in the Brownian case, it is sufficient to prove the formula for a measure η such that $\eta = \rho\left(-\delta^{V^\nu} v\right)$ with $v \in L^\infty(\nu, H)$. In that case we have

$$H(\eta|\nu) = E_\nu\left[-\delta^{V^\nu} v - \frac{|v|_H^2}{2}\right] = E_\nu\left[-\delta^{V^\eta} v + \frac{|v|_H^2}{2}\right]$$

Since $E_\nu\left[\delta^{V^\eta} v\right] = 0$, we get the result. □

Theorem VII.3. *Assume that there is a $\nu \in \mathcal{P}(\alpha, \beta)$, which is invertible in the sense of Definition VII.2. Then for any $a \in L_a^0(\nu, H)$ note (A, V^ν) the solution of (5.74) on $(W, \mathcal{F}^\nu, \nu)$ we defined in Lemma VII.2 . Further assume that $H(A\nu|\nu) < \infty$. We then have*

$$2H(A\nu|\nu) \leq E_\nu\left[|a|_H^2\right]$$

with equality if and only if

$$a + v \circ A = 0$$

where v is the Girsanov drift of $A\nu$ with respect to ν.

Proof: For convenience of notations we set $\eta := A\nu$. From Proposition VII.9, the finite entropy condition reads $E_\eta[|v|_H^2] < \infty$ so that

$$E_\nu[\delta^{V^\eta} v] = 0$$

Moreover $v \circ A \in L_a^2(\nu, H)$ so that

$$E_\nu\left[\delta^{V^\nu}(v \circ A)\right] = 0$$

Hence we obtain

$$\begin{aligned} 2H(\eta|\nu) &= E_\eta\left[|v|_H^2\right] \\ &= E_\eta\left[\delta^{V^\eta} v - \delta^{V^\nu} v\right] \\ &= -E_\eta\left[\delta^{V^\nu} v\right] \\ &= -E_\nu\left[\delta^{V^\nu \circ A} v \circ A\right] \end{aligned}$$

On the other hand from Lemma VII.2 $\nu-$a.s

$$A = U^\nu(V^\nu + a)$$

where U^ν is the $\nu-$inverse of V^ν. Hence

$$V^\nu \circ A = V^\nu + a$$

from which we obtain

$$\begin{aligned} 2H(\eta|\nu) &= -E_\nu\left[\delta^{V^\nu}(v\circ A)\right] - E_\nu\left[< a, v\circ A >_H\right] \\ &= -E_\nu\left[< a, v\circ A >_H\right] \end{aligned}$$

As in the Brownian case, Proposition VII.9 and the Cauchy-Schwartz inequality yield the result. □

9. Some further extensions

The hypothesis $(H1)$ is the good one to get a direct link between the invertibility of the shift, and the existence of a unique strong solution. However it is possible to weaken this hypothesis if we just want to get pathwise uniqueness among solutions of law $\nu \in \mathcal{P}(\alpha,\beta)$. In this section we introduce a new hypothesis : $(H2)$ will be weaker than $(H1)$. We say that a $\nu \in \mathcal{P}(\alpha,\beta)$ satisfies $(H2)$ if there exists a unique $V : W \to W$ such that (I_W, V) is a weak solution of (1.67) on $(W, \mathcal{F}^\nu, \nu)$. This means that under $(H1)$ there exists a unique Brownian motion $t \to V_t$ on $(W, \mathcal{F}^\nu_\cdot, \nu)$ such that $\nu-$a.s. we have

$$I_W = x + \int_0^\cdot \alpha_s dV_s + \int_0^\cdot \beta_s ds$$

where $t \to W_t$ still denotes the coordinate process. Under $(H1)$, we still call $V : W \to W$ the Brownian transform of ν. If $(H2)$ also holds, this definition is consistent with the previous one. Moreover, every result from Section 5 to 8 have been written so that it remains true if you exchange $(H1)$ with $(H2)$, and it is easy to see that the definition of the invertibility as well as Proposition VII.4 and Proposition VII.5 still hold under $(H2)$.

Definition VII.3. *Assume that $(H0)$ holds, let $\nu \in \mathcal{P}(\alpha,\beta)$ be such that $(H2)$ holds, and let V be the Brownian transform of ν. We set the following definitions :*

(i) We say that a solution (X,B) of (1.67) on a complete space $(\Omega,\mathcal{G},\mathbb{P})$ is a $\nu-$ solution of (1.67) if $X\mathbb{P} = \nu$ and $\mathbb{P}-$a.s. $V\circ X = B$

(ii) We say that a $\nu-$solution of (1.67) on $(W,\mathcal{F}^\mu,\mu)$ is a strong solution on the Wiener space, if for any W valued B defined on any complete space $(\Omega,\mathcal{G},\mathbb{P})$ such that $B\mathbb{P} = \mu$, we have that $(U(B),B)$ is a $\nu-$solution.

(iii) We say that (1.67) has a unique $\nu-$solution, if there is a strong $\nu-$solution U on the Wiener space such that for any other $\nu-$solution (X,B) defined on a complete space $(\Omega,\mathcal{G},\mathbb{P})$ we have $\mathbb{P}-$a.s. $X = U(B)$

The next proposition shows that there are lot of such solutions

Proposition VII.10. *Assume that $(H0)$ holds, let $\nu \in \mathcal{P}(\alpha,\beta)$ be such that $(H2)$ holds, and let V be the associated Brownian transform. Moreover, let $(\Omega,\mathcal{G},\mathbb{P})$ be a complete probability space. Then for any $X \in M_\mathbb{P}((\Omega,\mathcal{G}),(W,\mathcal{F}^\nu))$ such that $X\mathbb{P} = \nu$, $(X, V\circ X)$ is a $\nu-$solution to (1.67).*

Proof: Since $\widetilde{\mathbb{P}} - a.s.$ $\widehat{V}\circ\widetilde{U} = V\circ\pi^W\circ\widetilde{U} = V\circ U\circ\pi$, the result follows directly from Theorem VII.1 □

We now relate the $\nu-$invertibility to the notion of $\nu-$solutions :

Theorem VII.4. *Assume that $(H0)$ holds, let $\nu \in \mathcal{P}(\alpha,\beta)$ be such that $(H2)$ holds, and let V be the associated Brownian transform. Then we have*

(i) V is $\nu-$right invertible with inverse U if and only if (1.67) has a strong $\nu-$solution U on the Wiener space. Moreover when one of these conditions is filled, both U are the same

(ii) V is $\nu-$invertible with inverse U if and only if (1.67) has a unique $\nu-$solution U. Moreover when one of these conditions is filled, both U are the same

Proof: We first prove (i). Assume that U is a strong $\nu-$solution on the Wiener space. By definition we then have that $(U(I_W), I_W)$ is a $\nu-$solution. In particular $\mu - a.s.$ $V \circ U(I_W) = I_W$. This means that U is the right inverse of V. Conversely, assume that $\mu - a.s.$ $V \circ U = I_W$. Then $V \circ U(B) = B$ $\mathbb{P} - a.s.$ for any B such that $B\mathbb{P} = \mu$ on a space $(\Omega, \mathcal{G}, \mathbb{P})$. Hence $(U(B), B)$ is a $\nu-$solution. We now turn to the proof of (ii). Assume that (1.67) has a unique strong solution U. Since (I_W, V) is a $\nu-$solution it implies $\nu - a.s.$ $I_W = U(V)$. This prove that U is the left inverse of V. Since U is a strong $\nu-$solution the invertibility follows from (i). Conversely, assume that V is invertible with inverse U. Since it is right invertible (i) implies that for any B on a space $(\Omega, \mathcal{G}, \mathbb{P})$, $U(B)$ is a $\nu-$ solution. Let (X, B) be a $\nu-$solution on that space, we have $\mathbb{P} - a.s.$ $V \circ X = B$. By applying U to both sides, we get $\mathbb{P} - a.s.$ $X = U \circ V \circ X = U(B)$. $\square$

Of course the notion of $\nu-$solution does not fit in the applications in the general case, and to make our results relevant we have to find a condition under which the $\nu-$solutions are exactly the solutions with law ν : in that case, the last theorem would provide a condition for the pathwise uniqueness to hold among solutions of law ν.

Proposition VII.11. *Assume that $(H0)$ holds and let $\alpha \in \mathcal{P}(\alpha, \beta)$. Then, $(H2)$ holds for ν, and the $\nu-$solution are exactly the solutions of (1.67) with law ν if and only if, for any probability space $(\Omega, \mathcal{G}, \mathbb{P})$ and for any filtration $\mathcal{G}_.$ on that space we have :*

(i) For any (X, B) which is solution of (1.67) on that space we have $\mathcal{F}_.^B \subset \mathcal{F}_.^X$ where $\mathcal{F}_.^B$ (resp. $\mathcal{F}_.^X$) denotes the filtration generated by B (resp. by X).

(ii) For any $X \in \mathcal{R}(\mathbb{P}, \nu)$ which is adapted to $\mathcal{G}_.$, if there are two $(\mathcal{G}_.)-$Brownian motion B and $\widetilde{B}$ in $\mathcal{R}(\mathbb{P}, \mu)$ such that both (X, B) and $(X, \widetilde{B})$ are solutions of (1.67) on that space, with the filtration $\mathcal{G}_.$, then $\mathbb{P} - a.s.$

$$B = \widetilde{B}$$

(iii) For any $X \in \mathcal{R}(\mathbb{P}, \nu)$ there is a Brownian motion $B \in \mathcal{R}(\mathbb{P}, \mu)$ such that (X, B) is a solution of (1.67) on that space, with the filtration $\left(\mathcal{F}_t^X\right)$.

Proof: Assume that $(H2)$ holds for ν and that the $\nu-$solutions are exactly the solutions of law ν. Also note V the Brownian transform of ν. Let (X, B) be a solution, we have $\mathbb{P} - a.s.$ $V \circ X = B$. Together with $\mathcal{F}_.^V \subset \mathcal{F}_.^W$, this implies $\mathcal{F}_.^B \subset \mathcal{F}_.^X$ (for instance see the proof of Proposition VI.10). On the other hand, let $(\mathcal{G}_.)$ be a filtration on that space and let $X \in \mathcal{R}(\mathbb{P}, \nu)$ be $(\mathcal{G}_t)$ adapted. From Proposition VII.10 we have that $(X, V \circ X)$ is solution of (1.67) for the filtration $(\mathcal{F}_.^X)$, on the same space : this proves (iii). Finally, if (X, B) and $(X, \widetilde{B})$ are two solutions of law ν, from the definition of the $\nu-$solution we have $\mathbb{P} - a.s.$ $B = V \circ X = \widetilde{B}$ i.e. (ii) is proved. Conversely, we now assume that (i),(ii) and (iii) holds. By taking the space $(W, \mathcal{F}^\nu, \nu)$, since $I_W\nu = \nu$, there is a V such that (I_W, V) is a solution of (1.67) on $(W, \mathcal{F}^\nu, \nu)$ for the filtration generated by the coordinate process. Moreover (ii) implies this V is unique : we have $(H2)$ and V is the Brownian transform of ν. We now consider a solution (X, B) of (1.67) on a space $(\Omega, \mathcal{G}, \mathbb{P})$ for a filtration $(\mathcal{G}_t)$. From (i) we know that B is adapted to $\mathcal{F}_.^X \subset \mathcal{G}_.$ so that it is also a $\left(\mathcal{F}_t^X\right)$ $-$Brownian motion on the same space. Hence (X, B) is also a solution on that space for the filtration generated by $\left(\mathcal{F}_t^X\right)$. On the other hand Proposition VII.10 implies that $(X, V \circ X)$ is a solution of (1.67) for the filtration $\left(\mathcal{F}_t^X\right)$. Hence, by applying (ii) on the space $(\Omega, \mathcal{G}, \mathbb{P})$ for the filtration $\left(\mathcal{F}_t^X\right)$, we have $\mathbb{P} - a.s.$ $V \circ X = B$, i.e. (X, B) is a $\nu-$solution. $\square$

Part 3

Towards stochastic differential Geometry

VIII

Invertibility of adapted shifts on the Lie-Wiener space

ABSTRACT. Introduction : Lepingle's exponential. The left derivative on the Lie-Wiener space and the stochastic integral. Girsanov shift on the Lie-Wiener space. Invertibility of the Girsanov shift. Criterion of invertibility based on the entropy for the Lie-Wiener space. Remark on the invertibility of adapted perturbations of the identity on the space of the paths with values in a Riemanian manifold.

1. Introduction

In this chapter we study the transformations of measures induced by adapted shifts in the context of Lie groups (see **[23]** and references therein), and we transport the notion of invertibility of adapted shifts introduced in **[53]**. We provide an analogous version of the main result of **[46]**, and we still relate it to the existence of a unique strong solution for a related stochastic differential equation. As we told it in the introduction, the main feature of what we called stochastic invertibility is to involve some adapted isomorphisms of probability spaces which preserve the filtrations. In this chapter we will work on the space of the paths with values in a finite dimensional Lie group where another adapted isomorphism of probability spaces will be provided by Lepingle's exponential (its definition will be recalled later). The main idea of this chapter is to pullback the isomorphisms by Lepingle's exponential. One reason to work on Lie groups is that, together with Lepingle's isomorphism we can use the stochastic Campbell-Haussdorf formula which yields easily precise results. The structure of this chapter is the following. In section 2 we recall the main features of Malliavin calculus on a Lie group and recall the definition of the Lie-Sobolev derivative L along translations, and of the associated divergence $L^\star$ (a derivative along rotations can be defined in the same way, however we won't use it here). Although we slightly modify the definition to make clearer the fact that the results come of an isomorphism of probability spaces, our presentation is essentially the same as in **[51]** and **[45]**. We then define the stochastic integral as the restriction of $L^\star$ to adapted process in Section 3 : we could have defined it directly (similarly to **[10]** in the case of a Riemanian manifold for instance). However it seamed interesting to notice that it is still the restriction to adapted processes of the divergence associated with translations on the flat space. We then define the perturbations of the identity by using the stochastic Campbell-Haussdorf formula and we transport the most important results of the transformations of measure induced by adapted shifts on the flat space. In Section 4 we define an innovation process. In Section 5 we introduce the notion of invertibility for adapted shifts. In Section 6, we then relate invertibility to the existence of a unique strong solution for stochastic differential equations, and we extend the criterion of invertibility based on the relative entropy within this framework. We end this chapter with a remark which explains how the same procedure may be used to translate the notion of invertibility on the space of the paths with values in a Riemanian Manifold. For convenience of notations, we only investigated the case of probabilities equivalent to the Wiener measure. However our results may be extended to the case of absolutely continuous probabilities by means of the results of Part II.

2. Preliminaries and notation

Let G be a finite dimensional, connected, locally compact Lie group. We recall that the Lie algebra $\mathcal{G}$ of the group G may be seen as the tangent space T_eG to the manifold G at its identity element e. Moreover the right (resp. left) inner product on G will be denoted by $\hat{R}_g$ (resp. $\hat{L}_g$), and for any $g \in G$ we define

$$Ad\,(g) : X \in \mathcal{G} \to Ad\,(g)\,X := \hat{L}_g * \hat{R}_{g^{-1}} * X$$

so that for any left invariant vector field $X \in \mathcal{G}$ and any $x \in G$ we have

$$\left[[Ad\,(g)\,X]\left(f \circ \hat{R}_g\right)\right](x) = (Xf)\,(xg)$$

We assume $\mathcal{G}$ to be an Hilbert space with a scalar product we note $\langle .,. \rangle_{\mathcal{G}}$

2.1. Malliavin calculus on the Lie algebra $\mathcal{G}$.

Let $W = \mathcal{C}_0\left([0,1], \mathcal{G}\right)$ be the space of the paths on the Lie algebra, and let H be the associated Cameron-Martin space which is defined by

$$H = \left\{ h : [0,1] \to \mathcal{G} \middle| h_. = \int_0^. \dot{h}_s ds; \int_0^1 |\dot{h}_s|^2_{\mathcal{G}} ds < \infty \right\}$$

We recall that H is an Hilbert space for the scalar product $\langle h, k \rangle_H = \int_0^1 < \dot{h}_s, \dot{k}_s >_{\mathcal{G}} ds$ for any $h, k \in H$. We note i the injection $H \hookrightarrow W$ which is dense and continuous. By identifying H with its dual thanks to the Riesz representation, we note $i^\star$ the injection $W^\star \hookrightarrow H$, which is also dense and continuous.

$$W^\star \hookrightarrow^{i*} H \hookrightarrow^{i} \mathcal{G}$$

In the sequel we shall always identify $W^\star$ with a subset of H by means of this latter injection, and the notation $|l|_H$ will always mean $|i^\star(l)\,|_H$. Moreover $\langle .,. \rangle_{\mathcal{G}^\star,\mathcal{G}}$ will denote the duality bracket between W and $W^\star$. The triplet (W, H, i) is an abstract Wiener space (see **[25]**). Let $\mathcal{B}\,(W)$ be the Borelian sigma-field on W. It is well know that there is a unique Borelian measure μ on $(W, \mathcal{B}\,(W))$ such that for each $l \in W^\star$

$$\int \mu\,(dw) \exp\left(i \langle l, w \rangle_{W^\star, W}\right) = \exp - \frac{|l|_H^2}{2}$$

This measure is called the Wiener measure. $L^p\,(\mu)$ still denotes the space of equivalence classes which identifies the elements of $\mathcal{L}^p\,(\mu)$ which are equal $\mu - as$, and the spaces $\mathcal{L}^0$ and L^0 are defined as in Section 2 of Chapter VI. We will still note ∇ (resp. δ) the weak Sobolev derivative (resp. the associated divergence) on the path space (W, H, μ) : we have defined it in Chapter I. In the sequel $L^0_a\,(\mu, H)$ (resp. $L^2_a\,(\mu, H)$) still denotes the subset of the $u \in L^0\,(\mu, H)$ (resp. in $L^2\,(\mu, H)$) such that $t \to \dot{u}_t$ is adapted to $(\mathcal{F}_t)$, where $(\mathcal{F}_t)$ denotes the augmentation of the filtration generated by the coordinate process $t \in [0,1] \to W_t(\omega) := \omega(t) \in \mathcal{G}$ with respect to μ. We also set $\mathcal{F} := \mathcal{F}_1$, which is the completion of the Borelian sigma field with respect to μ, and the unique extention of μ to $\mathcal{F}$ will still be denoted by μ. We recall that $L^2_a\,(\mu, H) \subset Dom_2\,(\delta)$ and that

$$\delta u = \int_0^1 \dot{u}_s dW_s$$

We will still note δu the stochastic integral of $u \in L^0_a\,(\mu, H)$. Let $t \in [0,1]$, we note $\pi_t u$ the process u stopped at time t which is defined by

$$(\pi_t u)_. = \int_0^{t \wedge .} \dot{u}_s ds$$

Hence, with this notations

$$\delta \pi_t u = \int_0^t \dot{u}_s dW_s$$

If B is a Wiener process and if $u \in L^0\,(\mu, H)$ is integrable with respect to B, we will note

$$\delta^B \pi_t u = \int_0^t \dot{u}_s dB_s$$

and for that reason we will sometimes note $\delta^W u$ instead of δu when we will want to stress the fact that we are dealing with a stochastic integral with respect to the coordinate process $t \to W_t$.

2.2. Brownian motion, and Malliavin calculus on the Lie group G.
It might seem astonishing to introduce Malliavin calculus, while we are mainly interested in the study of "adapted shifts" on the Lie Group. As a matter of fact we will use it to get easily a stochastic integral form the divergence. We follow the construction of **[51]** and **[45]** and all the results of this subsection can be found with further details in the above references, up to a slight modification (see below). Let $C_G = C_e([0,1], G)$ where e is the identity element of G, and still note $(t,\omega) \in [0,1] \times W \to W_t \in \mathcal{G}$ the coordinate process on the space of the paths with values the Lie algebra, which is a Wiener process under μ. With the same notation as **[45]**, we note $p : \omega \in W \to p(\omega) \in C_G$ the unique strong solution of the equation

$$X_t = e + \int_0^t X_s dW_s \tag{2.75}$$

By definition it is such that for every smooth function $f : G \to \mathbb{R}$ with compact support we have $\mu - a.s.$ for any $t \in [0,1]$

$$f(p_t) = f(e) + \int_0^t (H_i f)(p_s)\, d^\star W_s^i \tag{2.76}$$

where the stochastic integral is written in the Stratonovich prescription (see **[22]**), where the (H^i) are a basis of the Lie algebra whose elements are identified to the associated derivatives, and where $t \to W_t$ is the coordinate process which is Brownian on $(W, \mathcal{F}^\mu, \mu)$. As we recalled it the exponential has the further property to be an isomorphism of filtered probability spaces. Indeed, the existence of the stochastic logarithm (see **[20]**) yields that p is invertible in the sense that there is a $q : C_G \to W$ such that $q\nu << \mu$ (these laws are the same) and $\mu-$ a.s.

$$q \circ p = I_W$$

and $\nu-$a.s.

$$p \circ q = I_G$$

We also recall that $t \to p_t$ generates the same sigma field as $t \to W_t$ (see **[20]**), and we note ν, the image of μ though the mapping p (i.e. $\nu = p\mu$). Let $(\mathcal{H}_t)_{t\in[0,1]}$ be the augmented filtration of the coordinate process on C_G with respect to ν, and $\mathcal{H} := \mathcal{H}_1$, we say that $(C_G, \mathcal{H}, \nu)$ is the classical Lie-Wiener space. We further recall that p (see **[20]**) can be used to define a left invariant (but not torsion free) connection Γ on G, and so an associated notion of martingale on G (precisely of Γ-martingale in the sense of **[33]**) by saying that (M_t) is a martingale on G if and only if it is the image through the exponential of a martingale on $(W, \mathcal{F}^\nu, \nu)$. In the flat Wiener space $(W, \mathcal{F}, \mu)$, the Malliavin derivative was built thanks to the translations, and to the Cameron-Martin theorem. Hence, a natural way to build the Malliavin derivative is to find an operation on C_G which corresponds to the translations, and then to find a correspondance to transport the Malliavin Calculus from the flat algebra to the curved space G. Let $h \in H$ and $\tau_h : W \to W$ be the translation defined by

$$\tau_h = I_{\mathcal{G}} + h$$

where $I_{\mathcal{G}}$ is the identity map on W (i.e. $I_{\mathcal{G}} : \omega \in W \to I_{\mathcal{G}}(\omega) := \omega \in W$). The proposition 7 of **[20]** which relates the martingale decomposition of a semi-martingale on $\mathcal{G}$ to the martingale decomposition of its exponential on G suggests that the analogous transformation $T_h : C_G \to C_G$, may be defined by

$$T_h := e(\theta h) I_G$$

where I_G is the identity map on C_G (i.e. $I_G : c \in C_G \to I_G(c) := c$), and where $e(\theta h)$ is defined pathwise to be the solution of the ordinary differential equation

$$\left(\frac{de(\theta h) \circ c}{dt}\right)(t) = [e(\theta h) \circ c](t)\, Adc_t \dot{h}_t$$

with $[e(\theta h)(c)](0) = e$, for any $c \in C_G$.

Proposition VIII.1. *For any $h \in H$, $\mu - a.s$*

$$p \circ \tau_h = T_h \circ p$$

Proof: The result may be seen as a direct application of the Campbell-Haussdorf formula of [**20**]. However, we can also prove it more suggestively from the proof of Theorem 13. 7. 5 of [**45**]. For any $h \in H$ we consider the stochastic equation

$$X_{\cdot} = e + \int_0^{\cdot} X_s d^\star W_s + \int_0^{\cdot} X_s \dot{h}_s ds \tag{2.77}$$

Since $p\mu = \nu$ we have $\mu - a.s$

$$\frac{de(\theta h)(p)}{dt}(t) = [e(\theta h)(p)](t)\, Ad(p_t)\, \dot{h}_t$$

Therefore the stochastic integration by part formula implies for any $t \in [0,1]$

$$\begin{aligned}
(T_h \circ p)_t &:= [e(\theta h)(p)](t)\, p_t \\
&= e + \int_0^t e(\theta h)_s \circ p\; p_s d^\star W_s + \int_0^t e(\theta h)_s \circ p Adp_s \dot{h}_s p_s ds \\
&= e + \int_0^t e(\theta h)_s \circ p\; p_s d^\star W_s + \int_0^t e(\theta h)_s \circ p\; p_s \dot{h}_s ds \\
&= e + \int_0^t (T_h \circ p)_s\, d^\star W_s + \int_0^t (T_h \circ p)_s\, \dot{h}_s ds
\end{aligned}$$

which means that $(T_h \circ p)$ solves equation (2.77) On the other hand

$$\left(\int_0^t p_s d^\star W_s\right) \circ \tau_h = \int_0^t p_s \circ \tau_h d^\star W_s + \int_0^t p_s \circ \tau_h \dot{h}_s ds \tag{2.78}$$

Together with the fact that (p_t) is the solution of equation (2.75), equation (2.78) implies that $p \circ \tau_h$ satisfies for any$t \in [0,1]$

$$p_t \circ \tau_h = e + \int_0^t p_s \circ \tau_h d^\star W_s + \int_0^t p_s \circ \tau_h \dot{h}_s ds$$

i.e. $(p \circ \tau_h)$ also solves equation (2.77). By unicity of the solution to (2.77) we get the result. $\square$

In particular Proposition VIII.1 directly implies that $\{T_h, h \in H\}$ is a commutative group of transformations of C_G, which preserve the $\nu-$equivalence class (i.e. $T_h\nu \sim \nu$ for any $h \in H$). Specifically, we have $\nu - a.s.$ $T_h \circ T_k = T_k \circ T_h = T_{h+k}$ and $T_h \circ T_{-h} = I_G$ for any $h, k \in H$. For that reason it is natural to start from that group to build the Malliavin calculus.

Definition VIII.1. *Let $F : C_G \to \mathbb{R}$, F is said to be cylindrical if there is a $n \in \mathbb{N}$, a f in the Schwartz space of the smooth rapidly decreasing functions $\mathcal{S}(\mathbb{R}^n)$, and a n-uplet $(t_1, ..., t_n) \in [0,1]^n$ such that $\nu - a.s.$*

$$F(c) = f(c_{t_1}, ..., c_{t_n})$$

We note $\mathcal{S}_\nu(\mathbb{R})$ the set of this real valued cylindrical functions

As in the flat case (the same proof as in [**45**] holds) a monotone class argument together with the convergence of martingales implies :

Proposition VIII.2. *The cylindrical functions on C_G are dense in $L^p(\nu)$ for every $p > 1$*

In [**45**] and [**51**] both a left and a right derivatives are built. However, in this paper we will only be interested by the left derivative which has nicer properties, and we recall its construction. In order to more be consistent we slightly modify the definitions of [**45**] and we set

Definition VIII.2. *For any $F : C_G \to \mathbb{R}$ which is cylindrical we define for any $h \in H$:*

$$L_h F = \frac{d}{d\lambda} F \circ T_{\lambda h}\big|_{\lambda=0}$$

right]

Although this definition avoid heavy calculus, the results and the proof we present until the end of this sections are nearly the same as in [**45**] and [**51**].

Proposition VIII.3. *For any $h \in H$ L_h is a closable operator on $L^2(\nu)$ and we have $\mu - a.s.$*

$$(L_h F) \circ p = \nabla_h (F \circ p)$$

on the extended domain of L_h

Proof: Let $f \in \mathcal{S}_\nu(\mathbb{R})$, then since $p\mu = \nu$, we also have from Proposition VIII.2 :

$$\begin{aligned}
\nabla_h (F \circ p) &= \frac{d}{d\lambda} (F \circ p) \circ \tau_{\lambda h}|_{\lambda=0} \\
&= \frac{d}{d\lambda} F \circ (p \circ \tau_{\lambda h})\,|_{\lambda=0} \\
&= \frac{d}{d\lambda} F \circ (T_{\lambda h} \circ p)\,|_{\lambda=0} \\
&= \frac{d}{d\lambda} F (T_{\lambda h}) \circ p|_{\lambda=0} \\
&= \frac{d}{d\lambda} F (T_{\lambda h})\,|_{\lambda=0} \circ p \\
&= (L_h F) \circ p
\end{aligned}$$

Hence we directly have $\mu - a.s.$ $\nabla_h (F \circ p) = (L_h F) \circ p$. Let (F_n) be a sequence of cylindrical functions with the property that $F_n \to 0$ in $L^p(\nu)$ and such that $(LF_n)_n$ is a Cauchy sequence in $L^p(\nu, H)$. Then, since $p\mu = \nu$, $(F_n \circ p)_n$ converges to 0 in $L^p(\mu)$, and $(\nabla (F_n \circ p))$ is Cauchy in $L^p(\mu, H)$. Since ∇ is closable, we have $\lim_{n\to\infty} \nabla_h (F_n \circ p) \to 0$ in $L^p(\mu)$ for any $h \in H$. Hence for every $h \in H$, $\lim_{n\to\infty} L_h F_n \to 0$ in $L^p(\nu)$ and the operator is closable. By taking the limit we have directly $(L_h F) \circ p = \nabla_h (F \circ p)$ on the extended domain. □

We set $Dom_p(L)$ to be the set of $F \in L^p(\nu)$ with the property that there is a sequence $(F_n)_{n\in\mathbb{N}} \subset \mathcal{S}_\nu(\mathbb{R})$ such that $F_n \to F$ in $L^p(\nu)$ and $(LF_n)_{n\in\mathbb{N}}$ is Cauchy in $L^p(\nu, H)$. Then $Dom_p(L)$ is the domain of L on which L is defined by

$$LF = \lim_{n\to\infty} LF_n$$

We call L the $Lie - Sobolev$ derivative. Note that the end of the last proof shows that $F \in Dom_p(L)$ if and only if $F \circ p \in Dom_p(\nabla)$.

Proposition VIII.4. *For any $u \in Dom_p(L)$, we have $\mu - a.s.$*

$$(L_u F) \circ p = \nabla_{u \circ p} (F \circ p)$$

Proof: Let (h_i) be an orthogonal basis of H, then :

$$\begin{aligned}
(L_u F) \circ p &= \left(\sum_i L_{h_i} F \langle h_i, u \rangle_H \right) \circ p \\
&= \sum_i (L_{h_i} F) \circ p \langle h_i, u \circ p \rangle_H \\
&= \sum_i \nabla_{h_i} (F \circ p) \langle h_i, u \circ p \rangle_H \\
&= \langle u \circ p, \nabla (F \circ p) \rangle_H \\
&= \nabla_{u \circ p} (F \circ p)
\end{aligned}$$

which is the result □

Since $L : F \in \mathcal{S}_\nu(\mathbb{R}) \subset L^p(\mathbb{R}) \to LF \in L^p(\nu, H)$ has a dense support, and its adjoint $L^\star : Dom_p(L^\star) \to L^p(\nu)$ is well defined. We note $Dom_p(L^\star)$ the set of the random variables $\eta \in L^p(\nu, H)$, such that for any $\phi \in Dom_p(L)$ (where $\frac{1}{p} + \frac{1}{q} = 1$). $E_\mu \left[\langle \nabla \phi, \eta \rangle_H \right] \leq c_{p,q} \left(|\phi|_{L^q(\nu)} \right)$. For any $\eta \in Dom_p(L^\star)$, $L^\star \eta$ is characterised by the relation $E_\mu [\phi L^\star \eta] = E_\mu \left[\langle L\phi, \eta \rangle_H \right]$, which holds for any $\phi \in Dom_p(L, E)$. It is easy to see that $\eta \in Dom_p(L^\star)$ if and only if $\eta \circ p \in Dom_p(\delta)$.

Proposition VIII.5. *Moreover $L^\star$ is characterised by the following relationship which holds for any $u \in Dom_p(L^\star)$ $\mu - a.s.$*

$$(L^\star u) \circ p = \delta (u \circ p)$$

Proof: By definition we have :

$$\begin{aligned}
E_\nu [L_u F] &= E_\nu [F L^\star u] \\
&= E_\mu [F \circ p (L^\star u) \circ p]
\end{aligned}$$

On the other hand

$$\begin{aligned}
E_\nu [L_u F] &= E_\mu [(L_u F) \circ p] \\
&= E_\mu [\nabla_{u \circ p} (F \circ p)] \\
&= E_\mu [\delta (u \circ p) F \circ p]
\end{aligned}$$

Since $p.$ generate the same sigma algebra as the coordinate process on W, we have $\mu - a.s.$

$$\delta (u \circ p) = (L^\star u) \circ p$$

□

3. Transformations of measure induced by adapted shifts on Lie groups

As we said it above, a natural approach is to start from the relationship between the martingale decomposition on C_G and the martingale decomposition on W which is given by the the proposition 7 of **[20]**, and then to use a Cambpell-Haussorf type formula as a rule of translation. Before we start this work it is important to notice that for any $u \in \mathcal{L}_a^0(\nu, H)$, since $p\mu = \nu$ and by unicity of solution to exponential equations of **[20]** the equivalence class of $p(I_\mathcal{G} + u \circ p) \in L^0(\mu, H)$ only depends on the equivalence class of $u \in L_a^0(\nu, H)$.

Definition VIII.3. *Let $u \in L_a^0(\nu, H)$, and*

$$\Omega = \{c \in C_G | u \in H\}$$

Let $\tau_u : W \to W$ be the translation of the path of the Lie algebra along the drift $u \circ p$, which is defined by

$$\tau_u = I_{\mathcal{G}} + u \circ p$$

where $I_{\mathcal{G}}$ is the identity map on W. And we define $T_u : C_G \to C_G$ by

$$T_u := e(\theta u) I_G$$

where I_G is the identity map on C_G, and where $e(\theta u) : (s, c) \in [0,1] \times C_G \to e(\theta u)_s(c)$ is defined pathwise to be the solution of the ordinary differential equation

$$\left(\frac{de(\theta u)(c)}{ds}\right)(s) = [e(\theta u)(c)](t) Ad(c_t) \dot{u}_t \circ c$$

with the initial condition $e(\theta u)(c)_0 = e$ if $c \in \Omega$, and by $e(\theta u)(c) = e$ if $c \in \Omega^c$. We will say that τ_u (resp. T_u) is the perturbation of the identity associated with the shift u on W (resp. on C_G).

The Proposition VIII.6 is a generalization of the Proposition VIII.1. As in this latter case, although the result could be obtained from the Campell-Haussdorff formula which is given in the Proposition 5 of **[20]**, we prefer to prove it in a more suggesting way.

Proposition VIII.6. *For any $u \in L_a^0(\nu, H)$ we have $\mu - a.s.$*

$$p \circ \tau_u = T_u \circ p$$

where T_u and τ_u are given by Definition VIII.3

Proof: Since $p\mu = \nu$ we have $\mu - a.s$

$$\frac{de(\theta u)(p)}{dt}(t) = [e(\theta u)(p)](t) Ad(p_t) \dot{u}_t \circ p$$

Hence using the stochastic integration by part formula, we have :

$$\begin{aligned}(T_u \circ p)_t &= e + \int_0^t e(\theta u)_s \circ p\, p_s d^\star W_s + \int_0^t e(\theta u)_s \circ p A d p_s \dot{u}_s \circ p\, p_s ds \\ &= e + \int_0^t e(\theta u)_s \circ p\, p_s d^\star W_s + \int_0^t e(\theta u)_s \circ p\, p_s \dot{u}_s \circ p ds\end{aligned}$$

On the other hand, since $t \to u_t \circ p$ is adapted, we have :

$$\left(\int_0^t p_s d^\star W_s\right) \circ \tau_u = \int_0^t p_s \circ \tau_u d^\star W_s + \int_0^t p_s \circ \tau_u \dot{u}_s \circ p ds$$

so that $p_t \circ \tau_u$ also solves the stochastic differential equation :

$$p_t \circ \tau_u = e + \int_0^t p_s \circ \tau_u d^\star W_s + \int_0^t p_s \circ \tau_u \dot{u}_s \circ p ds$$

by unicity of the solution we get the result. □

From Proposition VIII.6, it is straightforward to see that, just as in the flat case, the set of the invertible API which induce a measure equivalent to ν is a non-commutative group. The notations of the next definition will be used implicitly in the sequel :

Definition VIII.4. *Let $u \in L_a^0(\nu, H)$ and τ_u and T_u be as in Definition VIII.3. We note $(\mathcal{F}^u)_{t\in[0,1]}$ the filtration generated by $\tau_u : (t, \omega) \in [0,1] \times W \to \tau_u(t, \omega)$ augmented with respect to μ, and $\mathcal{H}^u_{t\in[0,1]}$ the filtration generated by $T_u : (t, c) \in [0,1] \times C_G \to T_u(t, c) \in G$, augmeted with respect to ν*

Proposition VIII.7. *Let $u \in L_a^0(\nu, H)$ and let $(\mathcal{F}^u)_{t\in[0,1]}$, and $\mathcal{H}^u_{t\in[0,1]}$ be given by Definition VIII.4. Then we have $\mu - a.s.$*

$$E_\nu\left[\phi|\mathcal{H}_t^u\right]\circ p = E_\mu\left[\phi\circ p|\mathcal{F}_t^u\right]$$

for any $\phi \in C_b(C_G)$

Proof: For every $f \in \mathcal{S}(\mathbb{R}^n)$ we have :

$$\begin{aligned}
E_\mu\left[\phi\circ pf\left((p\circ\tau_u)_{t_1},...,(p\circ\tau_u)_{t_n}\right)\right] &= E_\mu\left[\phi\circ pf\left((T_u\circ p)_{t_1},...,(T_u\circ p)_{t_n}\right)\right]\\
&= E_\nu\left[\phi f\left((T_u)_{t_1},...,(T_u)_{t_n}\right)\right]\\
&= E_\nu\left[E_\nu\left[\phi|\mathcal{H}_t^u\right] f\left((T_u)_{t_1},...,(T_u)_{t_n}\right)\right]\\
&= E_\mu\left[E_\nu\left[\phi|\mathcal{H}_t^u\right]\circ pf\left((T_u\circ p)_{t_1},...,(T_u\circ p)_{t_n}\right)\right]\\
&= E_\mu\left[E_\nu\left[\phi|\mathcal{H}_t^u\right]\circ pf\left((p\circ\tau_u)_{t_1},...,(p\circ\tau_u)_{t_n}\right)\right]
\end{aligned}$$

Since τ_u and $p\circ\tau_u$ generate the same sigma algebra, we get the result. □

Let $u \in L_a^2(\nu, H)$, then $u\circ p \in L_a^2(\mu, H) \subset Dom_2(\delta)$ so that $u \in Dom_2(L^\star)$, and $(L^\star u)\circ p = \delta^W(u\circ p)$. Assume now that $u \in L_a^0(\nu, H)$, and let (τ_n) be an increasing sequence of stopping times on $(C_G, \mathcal{H}_1, (\mathcal{H}_t))$ such that for any $n \in \mathbb{N}$, $\pi_{\tau_n} u \in L_a^2(\nu, H)$ and $\nu - a.s.$ $\lim_{n\to\infty}\tau_n \to 1$. Then

$$\begin{aligned}
\nu\left(|L^\star\pi_{\tau_n}u - L^\star\pi_{\tau_m}u| > \epsilon\right) &= \mu\left(|\left(L^\star\pi_{\tau_n}u\right)\circ p - \left(L^\star\pi_{\tau_m}u\right)\circ p| > \epsilon\right)\\
&= \mu\left(|\delta^W\left((\pi_{\tau_n}u)\circ p\right) - \delta^W\left((\pi_{\tau_n}u)\circ p\right)| > \epsilon\right)\\
&= \mu\left(|\delta^W\left(\pi_{\tau_n\circ p}(u\circ p)\right) - \delta^W\left(\pi_{\tau_n\circ p}(u\circ p)\right)| > \epsilon\right)
\end{aligned}$$

On the other hand let $\sigma_n := \tau_n\circ p$ for every $n \in \mathbb{N}$, then $E_\mu\left[|\pi_{\sigma_n}u\circ p|_H^2\right] = E_\nu\left[|\pi_{\tau_n}u|_H^2\right] < \infty$ and σ_n is increasing and $\mu - a.s.$ converges to 1 . Hence it is well known that independantly of the choice of (τ_n), $\left(\delta^W\left(\pi_{\tau_n\circ p}(u\circ p)\right)\right)_n$ converges in $L^0(\mu)$ to a random variable which is noted $\delta^W(u\circ p)$. Hence we also have $\nu\left(|L^\star\pi_{\tau_n}u - L^\star\pi_{\tau_m}u| > \epsilon\right) \to 0$, $n, m \to \infty$ and $\left(L^\star\pi_{\tau_n}u\right)_n$ converges in $L^0(\nu)$independantly of the choice of τ_n. We note $L^\star u$ this limit, and we call it the stochastic integral of $u \in L_a^0(\nu, H)$ with respect to the coordinate process under ν. By construction, we still have :

$$(L^\star u)\circ p = \delta^W(u\circ p)$$

Note that we could have defined it directly, as it is sometimes made in the literature, but we wanted to show that it could still be seen the divergence associated with a derivative along translations. This relationship enables us to compose the stochastic integral with a shift on the lie group. Indeed, from Proposition VIII.6 we have $\mu - a.s.$

$$\begin{aligned}
(L^\star v)\circ T_u\circ p &= (L^\star v)\circ p\circ\tau_u\\
&= \delta^W(v\circ p)\circ\tau_u\\
&= \delta^W(v\circ p\circ\tau_u) + \langle v\circ p\circ\tau_u, u\circ p\rangle_H\\
&= \delta^W(v\circ T_u\circ p) + \langle v\circ T_u\circ p, u\circ p\rangle_H\\
&= \left(L^\star(v\circ T_u) + \langle v\circ T_u, u\rangle_H\right)\circ p
\end{aligned}$$

Since $p\mu = \nu$, we have proved :

Proposition VIII.8. *For any $u, v \in L_a^0(\nu, H)$ the following relationship holds $\nu - a.s.$*

$$(L^\star v)\circ T_u = L^\star(v\circ T_u) + \langle v\circ T_u, u\rangle_H$$

where T_u and T_v are given by Definition VIII.3

We can check quickly that, as it is expected, the process $t \to L^\star \pi_t \alpha$ has the following property of a stochastic integral :

Proposition VIII.9. *For any $\alpha \in L_a^0(\nu, H)$, $t \to L^\star \pi_t \alpha$ is a local martingale on $\left(C_G, (\mathcal{H}_t)_{t\in[0,1]}, \nu\right)$ with an associated increasing process $\langle L^\star \pi_. \alpha, L^\star \pi_. \alpha\rangle_t = |\pi_t \alpha|_H^2$.*

Proof: Let $\alpha \in L_a^2(\nu, H)$, from Proposition VIII.6 , Proposition VIII.7 , and from the property of the stochastic integral $\delta^W \pi_.$ on the Lie algebra, $\mu - a.s.$ we get :

$$\begin{aligned} E_\nu\left[L^\star \pi_t \alpha | \mathcal{H}_s\right] \circ p &= E_\mu\left[(L^\star \pi_t \alpha) \circ p | \mathcal{F}_s\right] \\ &= E_\mu\left[\left(\delta^W \pi_t \alpha \circ p\right) | \mathcal{F}_s\right] \\ &= \delta^W(\pi_s \alpha \circ p) \\ &= (L^\star \pi_s \alpha) \circ p \end{aligned}$$

Which shows that $t \to L^\star \pi_t \alpha$ is a martingale on $(C_G, \mathcal{H}_1, (\mathcal{H}_t), \nu)$. Similary

$$\begin{aligned} E_\nu\left[(L^\star \pi_t \alpha)^2 - |\pi_t \alpha|_H^2 | \mathcal{H}_s\right] \circ p &= E_\mu\left[(L^\star \pi_t \alpha \circ p)^2 - |\pi_t \alpha \circ p|_H^2 | \mathcal{F}_s\right] \\ &= E_\mu\left[\left(\delta^W \pi_t \alpha \circ p\right)^2 - |\pi_t \alpha \circ p|_H^2 | \mathcal{F}_s\right] \\ &= \left(\delta^W \pi_s \alpha \circ p\right)^2 - |\pi_s \alpha \circ p|_H^2 \\ &= \left[(L^\star \pi_s \alpha)^2 - |\pi_s \alpha|_H^2\right] \circ p \end{aligned}$$

and we get the increasing process. The generalization to $L_a^0(\nu, H)$ is the same as in the flat case. □

In the sequel we will also need a version of the martingale representation theorem :

Theorem VIII.1. *Let $L \in L^2(\nu)$, then there is a $\alpha \in L_a^2(\nu, H)$ such that*

$$L - E_\nu[L] = L^\star \alpha$$

And for any local martingale M on $\left(C_G, (\mathcal{H}_t)_{t\in[0,1]}, \nu\right)$, there is a $\alpha \in L_a^0(\nu, H)$ such that

$$M_t = M_0 + L^\star \pi_t \alpha$$

Proof: Let $L \in L^2(\nu)$, then $L \circ p \in L^2(\mu)$ and the representation theorem applies on the flat space that there is a $\beta \in L_a^2(\mu, H)$ such that

$$L \circ p = E_\mu[L \circ p] + \delta(\beta)$$

Moreover, since p generates the same sigma algebra as the coordinate process we know that there is a $\alpha \in L_a^2(\nu, H)$ such that $\beta = \alpha \circ p$. Hence we have $\mu - a.s.$

$$\begin{aligned} L \circ p &= E_\mu[L \circ p] + \delta(\alpha \circ p) \\ &= E_\nu[L] + (L^\star \alpha) \circ p \end{aligned}$$

Since $p\mu = \nu$ we have $\nu - a.s.$

$$L = E_\nu[L] + L^\star \alpha$$

The proof of the last claim can be deduced easily, either from the last result, or from the flat case by reapeating the same procedure. □

For any $u \in L_a^0(\mu, H)$, we adopt the same notations as **[47]**, and we note

$$\rho(-\delta u) := \exp\left(-\delta^W u - \frac{|u|_H^2}{2}\right)$$

Moreover, for any $u \in L_a^0(\nu, H)$ we also set

$$\rho(-L^\star u) := \exp\left(-L^\star u - \frac{|u|_H^2}{2}\right)$$

Theorem VIII.2. *Let $u \in L_a^0(\nu, H)$, and let T_u be as in Definition VIII.3. Then we have*

$$T_u \nu << \nu$$

Moreover, if we further assume that

$$E_\nu[\rho(-L^\star u)] = 1$$

then we also have

$$T_u \nu \sim \nu$$

Proof: If $u \in L_a^0(\mu, H)$, then $u \circ p \in L_a^0(\mu, H)$ so that we know from Theorem 2. 3. 1. of **[50]** that $\tau_u \mu << \mu$ so that $p\tau_u \mu << p\mu = \nu$. Together with Proposition VIII.6 it implies $T_u \nu = T_u(p\mu) = (T_u \circ p)\mu = (p \circ \tau_u)\mu = p(\tau_u \mu) << \nu$. Suppose now that $E_\nu[\rho(-L^\star u)] = 1$, then since

$$\begin{aligned} \exp\left(-L^\star u - \frac{|u|_H^2}{2}\right) \circ p &= \exp\left(-(L^\star u) \circ p - \frac{|u \circ p|_H^2}{2}\right) \\ &= \exp\left(-\delta^W(u \circ p) - \frac{|u \circ p|_H^2}{2}\right) \\ &= \rho\left(-\delta^W(u \circ p)\right) \end{aligned}$$

we have $E_\mu\left[\rho\left(-\delta^W(u \circ p)\right)\right] = 1$, and from Proposition 2. 3. 1. of **[50]**, we have $\tau_u \mu \sim \mu$ so that $p\tau_u \mu \sim p\mu = \nu$. Hence $T_u \nu = T_u \circ p\mu = p \circ \tau_u \mu \sim \nu$ □

We now give a sharp Girsanov theorem on the Lie group :

Theorem VIII.3. *Let $\widetilde{\nu}$ be a probability absolutely continuous with respect to ν. Then there is a $v \in L_a^0(\widetilde{\nu}, H)$ such that $(t, \omega) \to \tau_v(\omega, t)$ is a $(\mathcal{F}_t)$ −Brownian motion on $(W, \mathcal{F}, \widetilde{\mu})$, and $T_v \widetilde{\nu} = \nu$, (where τ_v and T_v are given by Definition VIII.3). Moreover*

$$\frac{d\widetilde{\mu}}{d\mu} = \rho\left(-\delta^W(v \circ p)\right)$$

$\widetilde{\mu} - a.s.$ where $\widetilde{\mu}$ is defined by $\frac{d\widetilde{\mu}}{d\mu} = \frac{d\widetilde{\nu}}{d\nu} \circ p$ Moreover $\widetilde{\nu} \sim \nu$ if and only if there is a $v \in L_a^0(\nu, H)$ such that $\nu - a.s.$

$$\frac{d\widetilde{\nu}}{d\nu} = \rho(-L^\star v)$$

v will be called the Girsanov shift associated with $\widetilde{\nu}$.

Proof: From the Theorem VIII.1 we know that there is a $\alpha \in L_a^0(\nu, H)$ such that $\nu - a.s.$

$$L_t = 1 + L^\star \pi_t \alpha$$

where $L_t = E_\nu\left[\frac{d\widetilde{\nu}}{d\nu}|\mathcal{H}_t\right]$. Hence we have $\mu - a.s.$

$$\begin{aligned} L_1 \circ p &= 1 + (L^\star \alpha) \circ p \\ &= 1 + \delta(\alpha \circ p) \end{aligned}$$

Let $\widetilde{\mu}$ be defined by $\frac{d\widetilde{\mu}}{d\mu} = \frac{d\widetilde{\nu}}{d\nu} \circ p$ so that $p\widetilde{\mu} = \widetilde{\nu}$. From Proposition VIII.7 (with the same notations) we have $E_\mu\left[\frac{d\widetilde{\mu}}{d\mu}|\mathcal{F}_t\right] = L_t \circ p$ and by applying Doob's optionnal stopping time theorem, it is easy to see that if we set $\tau_0 := \inf(\{t : L_t \circ p = 0\}) \wedge 1$, then $\widetilde{\mu} - a.s.$ we have $\tau_0 = 1$. Hence $\widetilde{\mu} - a.s.$ we have

$$L_1 \circ p = 1 + \int_0^t \frac{\dot{\alpha}_s \circ p}{L_s \circ p} L_s \circ p dW_s$$

where the stochastic integral has to be seen as the stochastic integral with respect to the semi-martingale $s \to W_s$ on $(W, \mathcal{F}, (\mathcal{F}_t), \widetilde{\mu})$. Then we define $v \in L^0_a(\widetilde{\nu}, H)$ to be $v_. = -\int_0^{\cdot} \frac{\dot{\alpha}_s}{L_s}$. Hence we have

$$L_t \circ p = 1 - \int_0^t \dot{v}_s \circ p L_s \circ p dW_s$$

we can integrate this stochastic differential equation with respect to the semi-martingale $W_.$ on $(W, \widetilde{\mu})$ and we get $\widetilde{\mu} - a.s.$

$$\frac{d\widetilde{\mu}}{d\mu} = L_1 \circ p = \exp\left(-\delta(v \circ p) - \frac{|v \circ p|_H^2}{2}\right)$$

Note that if $\widetilde{\nu} \sim \nu$, $L^\star v$ is well defined and, from Proposition VIII.6, we have $\widetilde{\mu} - a.s.$

$$\frac{d\widetilde{\nu}}{d\nu} \circ p = \exp\left(L^\star v - \frac{|v|_H^2}{2}\right) \circ p$$

so that $\widetilde{\nu} - a.s.$

$$\frac{d\widetilde{\nu}}{d\nu} = \rho(-L^\star v)$$

The converse is obvious. It is then easy to see that T_v is a Brownian motion on $(W, \mathcal{H}_., \widetilde{\nu})$. Indeed, since $\omega \to \tau_v(\omega)$ is a $(\mathcal{F}_t)-$Brownian motion on $(W, \mathcal{F}, \widetilde{\nu})$ and $T_v \circ p = p \circ \tau_v$, we get $T_v \widetilde{\nu} = T_v(p\widetilde{\mu}) = T_v \circ p\widetilde{\mu} = (p \circ \tau_v)\widetilde{\mu} = p(\tau_v \widetilde{\mu}) = p\mu = \nu$ □

As a matter of fact the integrability of v is related to the entropy and we have :

Proposition VIII.10. *Let $\widetilde{\nu}$ be a probability which is equivalent to ν and let v denote the Girsanov shift of $\widetilde{\nu}$. Then*

$$E_\nu\left[|v|_H^2\right] = 2H(\widetilde{\nu}|\nu)$$

Proof: Let $\widetilde{\mu}$ be as in Theorem VIII.3, then τ_v is a $(\mathcal{F}_t)-$Brownian motion on $(W, \mathcal{F}, \widetilde{\mu})$ and from lemma 2. 6 of **[15]** and Proposition 2. 11 of **[14]** , we have $2H(\widetilde{\mu}|\mu) = E_\mu\left[|v \circ p|_H^2\right]$. Hence we get

$$2H(\widetilde{\nu}|\nu) = 2H(p\widetilde{\mu}|p\mu) = 2H(\nu|\mu) = E_\mu\left[|v \circ p|_H^2\right] = E_\nu\left[|v|_H^2\right]$$

□

4. The innovation process

We first define the notion of representation of a measure by a shift on the Lie group, analogous to the definition of **[47]** in the flat case

Definition VIII.5. *We note*

$$\mathcal{R}_a(\nu, \widetilde{\nu}) = \left\{u \in L^2_a(\nu, H) : T_u \nu = \widetilde{\nu}\right\}$$

and we say that a $u \in L^0_a(\mu, H)$ represents $\widetilde{\nu}$ (with respect to ν) if $u \in \mathcal{R}_a(\nu, \widetilde{\nu})$

From Proposition VIII.6 $u \in \mathcal{R}_a(\nu, \widetilde{\nu})$ if and only if $\tau_u \mu = \widetilde{\mu}$ where $\widetilde{\mu}$ is given by Theorem VIII.3. Indeed $T_u \nu = T_u p\mu = p\tau_u \mu$ and $\widetilde{\nu} = p\widetilde{\mu}$. Hence $H(T_u \nu|\widetilde{\nu}) = H(p\tau_u \mu|p\widetilde{\mu}) = H(\tau_u \mu|\widetilde{\mu})$, so that $T_u \nu = \widetilde{\nu}$ if and only if $\tau_u \mu = \widetilde{\mu}$. In the sequel we will use this result directly.

Theorem VIII.4. *Let $\widetilde{\nu}$ be a probability absolutely continuous with respect to ν, and let v be the Girsanov shift associated with $\widetilde{\nu}$ (see Theorem VIII.3). Assume Furthermore that there is a $u \in \mathcal{R}(\nu, \widetilde{\nu})$ and set $R^u := \tau_v \circ \tau_u$. Then $(t, \omega) \to R^u_t(\omega)$ is a $(\mathcal{F}^u_t)-$ Brownian motion on $(W, (\mathcal{F}^u_.), \mu)$, where $(\mathcal{F}^u_t)$ is given by Definition VIII.4.*

Proof:

Let $\tilde{\mu}$ be as in Theorem VIII.3, then since $\tau_u\mu = \tilde{\mu}$, for any $f \in C_b(W)$ which is $\mathcal{F}_s$ measurable, we have :

$$\begin{aligned} E_\mu\left[\left((\tau_v \circ \tau_u)_t - (\tau_v \circ \tau_u)_s\right) f \circ \tau_u\right] &= E_{\tilde{\mu}}\left[\left((\tau_v)_t - (\tau_v)_s\right) f\right] \\ &= 0 \end{aligned}$$

Where the last equality is a consequence of Theorem VIII.3, from which we know that $(t,\omega) \to (\tau_v)_t(\omega)$ is a $(\mathcal{F}_t)$ −Brownian motion on $(W, \mathcal{F}, \mu)$. Hence $(t,\omega) \to (\tau_v \circ \tau_u)_t(\omega)$ is a martingale on on $(W, \mathcal{F}^u, \nu)$, and by Levy's criterion, it is also a Brownian motion with respect to $(\mathcal{F}^u_t)$. □

Definition VIII.6. *Let $u \in L^2(\nu, H)$ we note $\widehat{u}$ (resp. $\widehat{u \circ p}$) the dual predicable projection of u on $(\mathcal{H}^u_t)_{t\in[0,1]}$ (resp. on $(\mathcal{F}^u_t)_{t\in[0,1]}$) which is defined as the projection of u on the closed subspace of the elements of $L^2(\nu, H)$ (resp. of $L^2(\mu, H)$) which are adapted to $(\mathcal{H}^u_t)_{t\in[0,1]}$ (resp. to $(\mathcal{F}^u_t)_{t\in[0,1]}$) where $(\mathcal{H}^u_t)$ and $(\mathcal{F}^u_t)$ are as in Definition VIII.4*

For a $u \in L^0(\nu, H)$ we can note $\widehat{u \circ p}_t = \int_0^t E_\mu\left[\dot{u}_s \circ p|\mathcal{F}^u_s\right] ds$ and $\widehat{u}_t = \int_0^t E_\nu\left[\dot{u}_s|\mathcal{H}^u_s\right] ds$

Definition VIII.7. *Let $u \in L^2(\nu, H)$ and $\widehat{u}$ (resp. $\widehat{u \circ p}$) be the dual projection of u (resp. of $u \circ p$) on $(\mathcal{H}^u_t)_{t\in[0,1]}$ (resp. on $(\mathcal{F}^u_t)_{t\in[0,1]}$). Then the innovation of u (resp. of $u \circ p$) noted Z^u (resp. $Z^{u\circ p}$) is defined by $Z^u = T_{u-\widehat{u}}$ (resp. $Z^{u\circ p} = I_W + u \circ p - \widehat{u \circ p}$)*

The next proposition is well known (see **[50]** or **[47]** for some proofs).

Proposition VIII.11. *Let $u \in L^2_a(\nu, H)$, then $(t,\omega) \to Z^{u\circ p}_t(\omega)$ is a $(\mathcal{F}^u_t)$ −Brownian motion on $(W, \mathcal{F}^u_\cdot, \mu)$*

Note that together with Proposition VIII.6 and Proposition VIII.7, the last Proposition also implies that the innovation Z^u on the Lie-Wiener space is a Brownian motion on the lie group. However we won't use that result directly. In order to use the two last proposition, we need the following Proposition, which is an obvious (but useful) consequence of Proposition VIII.7 :

Proposition VIII.12. *For any $u \in L^2_a(\nu, H)$, we have $\mu - a.s.$*

$$\widehat{u} \circ p = \widehat{u \circ p}$$

where $\widehat{u}$ and $\widehat{u \circ p}$ are given by Definition VIII.7

In the flat case Corollary VIII.1 is a classical result widely used in stochastic mechanics, stochastic control, and also in information theory. Corollary VIII.1 is a Lie-Group version of it which was unknown.

Corollary VIII.1. *Let $\widetilde{\nu}$ be a probability such that $\widetilde{\nu} \sim \nu$ and $u \in \mathcal{R}_a(\nu, \widetilde{\nu})$ such that $u \in L^2_a(\nu, H)$ we note $\widehat{u}$ the dual projection of u on $(\mathcal{H}^u_t)_{t\in[0,1]}$. Then we have*

$$\hat{u} = -v \circ T_u$$

$\nu - a.s.$

Proof: Since $u \circ p \in L^2_a(\mu, H)$ we know from Proposition VIII.11 that $t \to Z^{u\circ p}_t$ is a Brownian motion with respect to $(\mathcal{F}^u_t)_{t\in[0,1]}$. On the other hand if we set $R^u = \tau_v \circ \tau_u = I_{\mathcal{G}} + u \circ p + v \circ p \circ \tau_u$ we know from Theorem VIII.4 is also a Wiener process with respect to the same filtration. On the other hand, from Definition VIII.7, we have :

$$R^u_t - Z^u_t = (v \circ p \circ \tau_u)_t + \widehat{u \circ p}_t$$

Hence $t \to R_t^u - Z_t^u$ vanish as a martingale with finite variation, and we have

$$v \circ p \circ \tau_u + \widehat{u \circ p} = 0$$

$\mu - a.s.$ From Proposition VIII.12 we have $\widehat{u \circ p} = \widehat{u} \circ p$. Moreover, we know from Proposition VIII.6 that $v \circ p \circ \tau_u = v \circ T_u \circ p$. Since $p\mu = \nu$ it implies that $\nu - a.s.$

$$v \circ T_u + \widehat{u} = 0$$

□

5. Invertibility

Definition VIII.8. *Let $u \in L_a^0(\nu, H)$,we say that T_u is invertible, with inverse T_v if and only if there is a $v \in L_a^0(\nu, H)$ which is the both sided almost sure inverse of T_u i. e if and only if $\nu - a.s$ $T_u \circ T_v = I_G$ and $\nu - a.s$ $T_v \circ T_u = I_G$, where T_u and T_v are given by Definition VIII.3.*

Proposition VIII.13. *Let $\widetilde{\nu}$ be a probability which is absolutely continuous with respect to ν, v be the associated Girsanov shift, and let $u \in \mathcal{R}_a(\nu, \widetilde{\nu})$. Then the following assertions are equivalent :*

(i) $T_u \circ T_v = I_G$ $\widetilde{\nu} - a.s.$
(ii) $T_v \circ T_u = I_G$ $\nu - a.s.$
(iii) $\tau_v \circ \tau_u = I_W$ $\mu - a.s.$

Proof: By iterating the Proposition VIII.6, we have $\mu - a.s.$ $p \circ \tau_v \circ \tau_u = T_v \circ T_u \circ p$. Since $p\mu = \nu$, we clearly have (iii) implies (ii). By taking the logarithm (see **[20]**) we get the converse implication. Hence (ii) and (iii) are equivalent. On the other hand

$$\begin{aligned}
\widetilde{\nu}(\{\omega | T_u \circ T_v = I_G\}) &= T_u\nu(\{\omega | T_u \circ T_v = I_G\}) \\
&= \nu(\{\omega | T_u \circ T_v \circ T_u = T_u\}) \\
&\geq \nu(\{\omega | T_v \circ T_u = I_G\}) \\
&= 1
\end{aligned}$$

Where the last inegality is a consequence of the inclusion $\{\omega | T_v \circ T_u = I_G\} \subset \{\omega | T_u \circ T_v \circ T_u = T_u\}$. Since $T_v\widetilde{\nu} = \nu$, symmetrically $\nu(\{\omega | T_v \circ T_u = I_G\}) \geq \widetilde{\nu}(\{\omega | T_u \circ T_v = I_G\})$ so that

$$\nu(\{\omega | T_v \circ T_u = I_G\}) = \widetilde{\nu}(\{\omega | T_u \circ T_v = I_G\})$$

which clearly implies the result. □

Proposition VIII.14. *Let $\widetilde{\nu}$ be a probability which is equivalent to ν, and let v be the associated Girsanov shift (see Theorem VIII.3). Further assume that there is a $u \in \mathcal{R}_a(\nu, \widetilde{\nu})$ such that $T_v \circ T_u = I_G$. Then we have $E_\nu[\rho(-L^\star u)] = 1$*

Proof: A straightforward calculation shows that

$$\begin{aligned}
\rho(-\delta(v \circ p)) \circ \tau_u \rho(-\delta(u \circ p)) &= \rho(-\delta(u \circ p + v \circ p \circ \tau_u)) \\
&= \rho(-\delta(\tau_v \circ \tau_u - I_{\mathcal{G}}))
\end{aligned}$$

Since $T_v \circ T_u = I_G$ $\nu - a.s.$, from Proposition VIII.6 we have $\tau_v \circ \tau_u = I_{\mathcal{G}}$ $\mu - a.s.$ so that $\rho(-\delta(v \circ p)) \circ \tau_u \rho(-\delta(u \circ p)) = 1$ Hence, from Theorem VIII.3 we have :

$$\begin{aligned}
E_\nu\left[\rho\left(-L^\star u\right)\right] &= E_\mu\left[\rho\left(-L^\star u\right)\circ p\right]\\
&= E_\mu\left[\rho\left(-\delta\left(u\circ p\right)\right)\right]\\
&= E_\mu\left[\frac{1}{\rho\left(-\delta\left(v\circ p\right)\right)\circ\tau_u}\right]\\
&= E_{\tau_u\mu}\left[\frac{1}{\rho\left(-\delta\left(v\circ p\right)\right)}\right]\\
&= E_{\widetilde{\mu}}\left[\frac{1}{\rho\left(-\delta\left(v\circ p\right)\right)}\right]\\
&= E_\mu\left[\frac{d\widetilde{\mu}}{d\mu}\frac{1}{\rho\left(-\delta\left(v\circ p\right)\right)}\right]\\
&= 1
\end{aligned}$$

□

6. Criterion of invertibility based on the relative entropy

We now consider the following equation

$$X = e + \int_0^\cdot X_t\left(dW_t - \dot{v}_t\circ X dt\right) \tag{6.79}$$

which means, that for any smooth f with compact support

$$f\left(X_t\right) = f\left(e\right) + \int_0^t \left(H_i f\right)\left(X_s\right)\star\left(dW_t - \dot{v}_t\circ X dt\right)$$

where $v\in L_a^0\left(\nu, H\right)$ is such that $E_\nu\left[\rho\left(-L^\star v\right)\right]=1$. Once again, we define two probabilities

$$\frac{d\widetilde{\mu}}{d\mu} = \rho\left(-\delta^W\left(v\circ p\right)\right)$$

and

$$\frac{d\widetilde{\nu}}{d\nu} = \rho\left(-L^\star v\right)$$

As above, we still note $\tau_v = I_W + v\circ p$, and T_v is defined as above. We also set $\widehat{T_v} = q\circ T_v$. From the definition of p, the Girsanov theorem directly implies

Proposition VIII.15. *(p,τ_v) is a weak solution of (6.79) on $\left(W, \mathcal{F}^{\widetilde{\mu}}_\cdot, \widetilde{\mu}\right)$ with law $\widetilde{\nu}$*

Proposition VIII.16. *Assume that there is a X defined on a space on a probability space $\left(\Omega, \mathcal{F}^P, P\right)$ such that $X\mathbb{P}=\widetilde{\nu}$. Then $\left(X, \widehat{T_v}\left(X\right)\right)$ is a solution of (6.79) on $\left(\Omega, \mathcal{F}^P, P\right)$ with the filtration $\left(\sigma\left(X_s, s\leq t\right)\right)_{t\in[0,1]}$*

Proof: Mutatis mutandis, the proof of Theorem VII.1 implies the result. With the suggesting notations of the introduction for the exponential it can be written shortly. We have

$$p = e + \int_0^\cdot p_t\left(d\left(\tau_v\right)_t - \dot{v}_t\circ p dt\right)$$

Since $X = p\circ\left(q\circ X\right)$ we get

$$X = e + \int_0^\cdot X_t\left(d\left(\left(\tau_v\right)_t\circ q\circ X\right) - \dot{v}_t\circ X dt\right)$$

On the other hand $\tau_v\circ q = q\circ\left(p\circ\tau_v\right)\circ q = q\circ T_v$, where we used that $p\circ\tau_v = T_v\circ p$. This yields the result.

□

Proposition VIII.17. *T_v is invertible if and only if (6.79) has a unique strong solution. In that case, let f denotes the measurable function which generates the solutions, and let T_u denote the inverse of T_v, then we have $\mu - a.s.$*

$$f = T_u \circ p$$

Proof: The fact that all the inverse images we will use are well defined can be checked exactly in the same way as we did it (with full details) in the Sections 3 and 4 of Chapter VII. We now begin the proof. Assume that T_v is invertible with inverse T_u we set $\widehat{T_u} = T_u \circ p$. If B is a Brownian motion on a space $\left(\Omega, \mathcal{F}^P, P\right)$ $\left(\widehat{T_u} \circ B\right) P = T_u \circ p\mu = T_u\nu = \widetilde{\nu}$. Since $\widehat{T_v} \circ \widehat{T_u} = q \circ (T_v \circ T_u) \circ p = I_W$ Proposition VIII.16 implies that $\left(\widehat{T_u} \circ B, B\right)$ is solution. Moreover, if we have a solution (X, B) on a probability space $\left(\Omega, \mathcal{F}^P, P\right)$ with a filtration $\mathcal{F}_t$, we have $X = \widehat{T_u} \circ \widehat{T_v} \circ X$. Since $\left(X, \widehat{T_v} \circ X\right)$ is solution on the same space with respect to the filtration generated by X and $\mathcal{F}_t^X \subset \mathcal{F}_t$ (the filtration generated by X), $B = \widehat{T_v} \circ X$. Hence $\widehat{T_u}(B) = \widehat{T_u} \circ \widehat{T_v} \circ X = X$. This proves that (6.79) has a unique strong solution. Conversely, assume that (6.79) has a unique strong solution, and note f the associated function. Since (p, τ_v) is a weak solution we have $p = f(\tau_v)$ i.e. $I_W = q \circ f \circ \tau_v$. On the other hand, (f, I_W) is also a solution on $(W, \mathcal{F}^\mu, \mu)$. Hence from Proposition VIII.16 $I_W = \widehat{T_v} \circ f = q \circ T_v \circ f = q \circ T_v \circ p \circ q \circ f = \tau_v \circ (q \circ f)$. This means that τ_v is invertible with inverse $\tau_u := qf$ i.e. $f = p \circ (q \circ f) = p \circ (\tau_u) = T_u \circ p = \widehat{T_u}$. But as we proved it above, the invertibility of τ_v is equivalent to the invertibility of T_v. □

The next theorem is a Lie-group version of the main result of **[47]**. This result has various applications to information theory and optimal transport. See **[47]** or even **[48]** for some these applications.

Theorem VIII.5. *Let $\widetilde{\nu}$ be a probability equivalent to ν, and let v be the Girsanov shift associated with $\widetilde{\nu}$ by Theorem VIII.3. Furthermore, assume that there is a $u \in \mathcal{R}_a(\nu, \widetilde{\nu})$, and let ϵ be the mean square error of the causal estimates (i.e. $\epsilon = E_\nu\left[|u - \widehat{u}|_H^2\right]$). Then we have :*

$$\epsilon = E_\nu\left[|u|_H^2\right] - 2H(\nu|\mu) \geq 0$$

Moreover the following conditions are equivalent :

- *T_v is invertible with inverse T_u (where T_v and T_u are given by Definition VIII.3)*
- $\epsilon = 0$
- $2H(\widetilde{\nu}|\nu) = E_\nu\left[|u|_H^2\right]$
- *The stochastic differential equation (6.79) has a unique strong solution which is given by $T_u \circ p$.*

Proof: From Proposition VIII.10 and Theorem VIII.1, we have $2H(\widetilde{\nu}|\nu) = E_\nu\left[|\widehat{u}|_H^2\right]$ Hence

$$\begin{aligned}
2H(\widetilde{\nu}|\nu) &= E_{\widetilde{\nu}}\left[|v|_H^2\right] \\
&= E_{T_u\nu}\left[|v|_H^2\right] \\
&= E_\nu\left[|v \circ T_u|_H^2\right] \\
&= E_\nu\left[|\widehat{u}|_H^2\right] \\
&\leq E_\nu\left[|u|_H^2\right]
\end{aligned}$$

with equality if and only if $\nu - a.s.$ $\widehat{u} = u$, and we also have

$$\begin{aligned}
\epsilon &= E_\nu\left[|u - \widehat{u}|_H^2\right] \\
&= E_\nu\left[|u|_H^2\right] - E_\nu\left[|\widehat{u}|_H^2\right] \\
&= \mathrm{E}_\nu\left[|u|_H^2\right] - 2H(\widetilde{\nu}|\nu)
\end{aligned}$$

moreover $\epsilon = 0$ if and only if $\widehat{u} = u \ \nu - a.s.$ and from Theorem VIII.1 this condition is also equivalent to $\nu - a.s.$, $u = -v \circ T_u$. Since $p\mu = \nu$ it is also equivalent to $u \circ p = -v \circ T_u \circ p \ \mu - a.s.$, and from Proposition VIII.6, it is also equivalent to $\mu - a.s.$ $u \circ p = -v \circ p \circ \tau_u$. Thus, we obtain

$$\begin{aligned} u \circ p + v \circ T_u \circ p &= u \circ p + v \circ p \circ \tau_u \\ &= I_{\mathcal{G}} + u \circ p + v \circ p \circ \tau_u - I_{\mathcal{G}} \\ &= \tau_u + v \circ p \circ \tau_u - I_{\mathcal{G}} \\ &= \tau_v \circ \tau_u - I_{\mathcal{G}} \end{aligned}$$

so that $\epsilon = 0$ if and only if $\mu - a.s.$ $\tau_v \circ \tau_u = I_{\mathcal{G}}$. From Proposition VIII.13 we know that this condition is equivalent to the invertibility of T_u. Proposition VIII.17 achieves the proof. □

Remark: Let $\left(W = C_0\left([0,1], \mathcal{R}^d\right), \mathcal{F}^\mu, (\mathcal{F}^\mu_t), \mu\right)$ be the classical Wiener space, let M be a d dimensional Riemanian Manifold (see [**23**]), and let W^M be the set of the continuous curves in M which start from a given $m_0 \in M$. It is well known that a Brownian motion on M can be built from the canonical Brownian motion $t \to W_t$ on W by the so called rolling without slipping procedure (for instance see [**22**], [**31**], or [**21**]). This procedure provides an isomorphism $\mathcal{I}$ of probability spaces between (W, μ) and $\left(W^M, \mu^M\right)$, where μ^M is the Wiener measure on M with the Riemanian connection. $\mathcal{I}$ also preserves the filtrations i.e. $\mathcal{I}_t$ generates the same filtration as $t \to W_t$. With other words, $\mathcal{I}$ is both adapted and invertible in the sense that there is a $\mathcal{I}^{-1}$ such that $\mathcal{I}^{-1}\mu^M << \mu$ and μ−a.s.

$$\mathcal{I}^{-1} \circ \mathcal{I} = I_W$$

and μ^Ma.s.

$$\mathcal{I} \circ \mathcal{I}^{-1} = I_M$$

where I_M is the identity map on W^M. $\mathcal{I}$ is called the Itô map and enables to develop processes on the flat space to W^M, while its inverse enables to get anti-developements. This ismorphism of filtered probability spaces enables to map directly the stochastic integral on W to a stochastic integral with respect to adapted vector fileds on W^M (see [**31**], or [**33**], [**9**] for an alternative approach). On the other hand, the same properties imply that quasi invariant flows on (W, μ) are canonically associated through $\mathcal{I}$ to quasi invariant flows on $\left(W^M, \mu^M\right)$. Hence the representations formula, as well as the variational formula for the relative entropy can be transported in a straightforward fashion on $\left(W^M, \mu^M\right)$ by using the Itô map instead of Lepingle's exponential. However, on a Riemanian manifold, we don't have the Campbel Hausdorf formula, and the calculus of the transformations analogous to our T_v may require to use the structural equations.

Bibliography

[1] Allinge, D., Mitter, S.K. (1980) *New results on the innovations problem for nonlinear filtering* Stochastics 4, no. 4, 339–348.

[2] Artstein, S., Ball, K. M. ,Barthe, F. and Naor (2004) *Solution of Shannon's problem on the monotonicity of entropy*, J. Amer. Math. Soc., 17(4):975–982.

[3] Boué, M. Dupuis, P. (1998) *A variational representation for certain functionals of Brownian Motion* The Annals of Probability, Vol 26, No. 4, 1641-1659

[4] Carlen, E.,Cordero-Erausquin, D. (2009) *Subadditivity of the entropy and its relation to Brascamp-Lieb type inequalities*, Geom. Funct. Analysis 19 , 373–405.

[5] Cattiaux, P., Leonard, C. (1994) *Minimization of the Kullback information of diffusion processes.*, Annales de l'I.H.P., Section B., Tome 30

[6] Dellacherie, C. and Meyer, P. A. (1975). *Probabilités et Potentiel Ch. I à IV*, Paris, Hermann.

[7] Dellacherie, C. and Meyer, P. A. (1980)*Probabilités et Potentiel Ch. V à VIII*, Paris, Hermann.

[8] Driver, B. K.(1993) *Toward calculus and geometry on path space*, Proceedings of symposia in Pure Mathematics, Volume 57

[9] Emery, M.(1989) *Stochastic calculus in Manifolds*, Springer-Verlag

[10] Fang, S. and Malliavin, P. (1993) *Stochastic Analysis on the Path Space of a Riemanian Manifold* J.F.A. 118, 249-274.

[11] Fernique, X. (2003) *Extension du théorème de Cameron- Martin aux translations aléatoires*, The Annals of Probability, Vol. 31, No. 3, 1296–1304.

[12] Feyel, D. and Üstünel, A. S.(2004) *Monge-Kantorovitch measure transportation and Monge-Ampère equation on Wiener space*, *Probab. Theor. Relat. Fields*, **128**, no. 3, pp. 347–385,

[13] Feyel, D. and Üstünel, A. S. (2004) *Solution of the Monge-Amp'ere Equation on Wiener Space for log-concave measures: General case*, Preprint Arxiv.org

[14] Föllmer, H. (1986) *Time reversal on Wiener space*, Lecture Notes in Mathematics

[15] Föllmer, H. (1986) *Time reversal of infinite-dimensional diffusions*, Stochastic Processes and their Applications Volume 22, Issue 1, 59–77.

[16] Föllmer, H.(1988) *Random fields and diffusion processes*, Ecole d' été de Saint Flour XV–XVII

[17] Föllmer, H. and Gandert, N. (1997) *Entropy minimization and Schrödinger process in infinite demensions*, The annals of probability vol 25, No 2, 901–926.

[18] Guerra, F. and Morato, L. (1983) *Quantization of dynamical systems and stochastic control theory*, Physical review D volume 27, Number 8

[19] Gross, L. (1975) *Logarithmic Sobolev inequalities*, Amer. J. Math.

[20] Hakim-Dowek, M., Lepingle, D. (1986) *L'exponentielle stochastique des groupes de Lie*, Sem. Proba. XX, p. 352–374

[21] Hsu, H.P. ,*Stochastic Analysis on Manifolds*, Graduate Studies in mathematics Volume 38.

[22] Ikeda, N. and Watanabe, S. (1981)

[23] Kobayashi, S. and Nomizu, K. (1963)*Foundations of Differential Geometry Vol I*, Wiley Classics Library. *Stochastic Differential Equations and Diffusion Processes*. North Holland, Amsterdam (Kodansha Ltd., Tokyo)

[24] Krylov,N.V. and Röckner, M.R. (2005) *Strong solutions of stochastic equations with singular time dependent drift*, Probab. Theory Related Fields, 131, no. 2, 154–196.

[25] Kuo, H. (1975) *Gaussian Measures in Banach Spaces*, Lect.Notes in Math., 463

[26] Lehec *A stochastic formula for the entropy and applications*, 2010 Preprint.

[27] Lassalle, R. (2012) *Invertibility of adapted perturbation of the identity on abstract Wiener space*, J. Funct. Anal. 262, No. 6, 2734-2776

[28] Lassalle, R. (2012) *Local invertibility of adapted shifts on Wiener space, under finite energy condition*, Stochastics and Stochastics Reports. DOI:10.1080/17442508.2012.72025

[29] Lassalle, R. (2012) *Local invertibility of adapted shifts on Wiener space and related topics* , Springer Proceedings in Mathematics & Statistics, Vol. 34, Malliavin Calculus and Stochastic Analysis : A Festschrift in Honor of David Nualart, ISBN 978-1-4614-5905-7.

[30] Malliavin, P. (1997). *Stochastic Analysis*, Springer Verlag.

[31] Malliavin,P.(1977) *Géométrie différentielle stochastique*, les presses universitaires de l'université de Montréal

[32] Mayer-Wolf, E. and Zakai, M. (2007) *Some relations between mutual information and estimation error in Wiener space*, The Annals of Appl. Proba. Vol. 7, No. 3, p. 1102–1116, 2007.

[33] Meyer, P.A. (1981)*Géométrie stochastique sans larmes*, Séminaire de probabilités de Strasbourg, 15 p. 44-102

[34] Mikami, T. (1999)*Dynamical systems in the variational formulation of the Fokker-Planck equation by the Wasserstein metric*, Series 464. Available at `http://eprints3.math.sci.hokudai.ac.jp/655/1/464.PDF`

[35] Mikami, T. (2002) *Optimal control for absolutely continuous stochastic processes and the mass transportation problem*, Elect.Comm.in. Probab. 7 199–213.

[36] Mikami ,T. and Thieullen,M. (2006) *Duality Theorem for Stochastic Optimal Control Problem*, Sto. Proc. Appl. 116,1815-1835 , MR230760

[37] Nelson, E. (1966) *Derivation of the Schrödinger equation from Newtonian mechanics* , Physical review vol.150, number 4.

[38] Nelson, E. (1985) *Quantum fluctuations*, Princeton series in physics.

[39] Nelson, E. (1967) *Dynamical theories of Brownian Motion*, Princeton university Press.

[40] Nualart, D. (1995). *The Malliavin calculus and related topics*, Prob. and Appl., 21, Springer.

[41] Sugita, H. (1988). *Positive generalized Wiener functions and potential theory over abstract Wiener spaces*, Osaka J. Math. Volume 25, Number 3 (1988), 665-696.

[42] Talagrand, M. (1996) *Transportation cost for Gaussian and other product measures*, Geom. Funct. Anal., 6, 587–600.

[43] Tsirelson , B.S. (1975) *An example of stochastic differential equation having no strong solution*, Theor. Prob. Appl. **20**, p. 416–418

[44] Üstünel, A. S. (1995) *Introduction to Analysis on Wiener Space*, Lecture Notes in Math. Vol. 1610. Springer.

[45] Üstünel, A. S. (2010) *Analysis on Wiener Space and Applications.*

[46] Üstünel, A. S. (2008) *A necessary and sufficient condition for the invertibility of adapted perturbations of identity on the Wiener space* , C.R. Acad. Sci. Paris, Ser. I, Vol.**346**, p. 897–900.

[47] Üstünel, A. S. (2009) *Entropy, invertibility and variational calculus of adapted shifts on Wiener space* , J. Funct. Anal. 257, no. 11, 3655–3689.

[48] Üstünel, A. S. (2010) *Persistence of invertibility on the Wiener space*, COSA, vol. 4, no. 2, p. 201–213.

[49] Üstünel, A. S. and Zakai, M. (1997) *The construction of filtrations on abstract Wiener space*, J. Funct. Anal. **143** , p. 10–32.

[50] Üstünel, A. S. and Zakai, M. (1999) *Transformation of Measure on Wiener Space*, Springer Verlag

[51] Üstünel, A. S. (1997) *Stochastic analysis on Lie Groups*, Stochastic Analysis and Related Topics VI, The Geilo Workshop, 1996, p. 129-158 (1997). Progress in Probability, Vol. 42

[52] Üstünel, A. S. (1987) *Representation of the distributions on Wiener space and stochastic calculus of variations* Journal of functional analysis, Volume 70, Issue 1.

[53] Üstünel, A. S. and Zakai, M. (2007) *Sufficient conditions for the invertibility of adapted perturbations of identity on the Wienerspace*. Probab. Theory Relat. Fields, **139**, p. 207-234.

[54] Veretennikov ,A. Yu. (1980) *On strong solutions and explicit formulas for solutions of stochastic integral equations*, Mat. Sb. (N.S.), Volume 111(153), Number 3, Pages 434–452

[55] Villani, C. (2003) *Topics in Optimal Transportation*. Graduate Series in Math., **58**. Amer. Math. Soc.

[56] Watanabe, S.(1987) *Analysis of Wiener Functionals (Malliavin Calculus) and its Applications to Heat Kernels*,The Annals of Probability Vol. 15, No. 1 (Jan., 1987), pp. 1-39

[57] Yan, J.A.(1987) *Developpement des distributions suivant les chaos de Wiener et applications a l'analyse stochastique*, Lecture Notes in Mathematics, 1987, Volume 1247/1987, 27-32

[58] Yasue, K.(1981) *Stochastic calculus of variations*, Journal of functional Analysis.

[59] Zakai, M. (2005) *On mutual information, likelihood ratios and estimation error for the additive Gaussian channel*, IEEE Trans. Inform. Theory, **51**, 3017–3024.

[60] Zakai, M. (1969) *On the optimal filtering of diffusion processes*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 11,p. 230–243.

[61] Zambrini, J.C. (1986) *Stochastic mechanics according to E. Schrödinger*, Physical review A , volume 33, number 3.

[62] Zambrini, J.C. (1986) *Variational processes and stochastic versions of mechanics*, J. math. Phys. 27 (9).